

```
QY 602 ACAGAAAGTAAATAATGGAAATTTGTCTAGGAGATCCAGGATGAAGTAGAAAAATCA 661
Db 486 ACGAGAGCCAAAATAATGGAATACTGTCTGGAGATTTCCAGGCATGAAGCAGAGAACCA 545
QY 662 GAAATCAAGTGAAGAACACAGATGCTTCGGGCAAAATAGAGAAATATAATGTTCCGCTGAAC 721
Db 546 GAGACGAGCGAANAACACAGAAACTTTCAGGCAAAATAGAGAAATACAAAGCTTCCACTGAAT 605
QY 722 AGGCTTAAGATGATGTTTGAAGAAAGGTGAACCAACTCAAACTAAGATTTCTCCGGGCCCAA 781
Db 606 AGACTGAAGATGATGTTTGAAGAAAGGTGAACCAACTCAAACTAAGATTTCTCCGGGCCCAA 665
QY 782 AGCCGAAGTCAAGTGAAGAAAGATCTCTGAAGAAAGCTATTCTCTAGATGACCTGGAA 841
Db 666 AGCCGAATGCGGGTGAAGAGGCTCTCTGAAGAAAGCTATTCTCTAGATGACCTGGAA 725
QY 842 ATAGGCCAGGTCAAGTGTCTATCTTCTACATTTTCACTCGGAGAAATAATGAGATGAGCGA 901
Db 726 ATAGTGGGGTCAATTTGTCTTCTCTGCAATTCAACTCGGAGAAATAATGAGATGAGCGG 785
QY 902 AATCTGAACCTTCCACGCCCTCTCAGAAACCTCTATAAAGGATCGAATGGCCAAAGTACCAG 961
Db 786 AATCTGAGCTGCCACGCCCTCTCAGAAACCTCCATAAAGGACCGCATGGCCAAAGTACCAG 845
QY 962 GCAGCTGTGTCCAAACAAAGCAGCTCAACCACTATAAAGGATCGAATGGCCAAAGTACCAG 1021
Db 846 GCTCAGTGTCCAAAGCAGCAGCCAGCCAGCTACACAAATGAGCTGAATAACCAAGT... 902
QY 1022 GCGAAATCAAAATTCATAAAATGAGCAAAAGAGAGATGTGCCCCAGCTCTCGAGTGC 1081
Db 903 ---GAAAGCAAAACTCATAAATGGGAAACAGAGAGAGATGTGCCCCAGCTCCGAGGCC 959
QY 1082 TGATCAACCATCAGGAAGGGGAAAGATTTCTGCAAAATCAGAAATAGCCTGGCAGTCCGT 1141
Db 960 TGCAGCTCCATCAGGAAGGAGCAAGGTTTCTACAACTGAGAAATAGTCTAGTGGCCCTC 1019
QY 1142 TCCACCCCTCGCGAAGATGACTCCCGTGAATCCGAGTTTAAGAGTGAAGTTCAACAGCCT 1201
Db 1020 TCTGTCCCTGTGAAGATGACACCTGTAATCCAGGTGAAGAGCGAGGCCAGCAGCCT 1079
QY 1202 GTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCCTCCAGTCTTCTGAAAGTTCTCCT 1261
Db 1080 ATGACCCCTAAGCCGCTGAGTCTCTAGCCAGCAGACCCCACTCTTCGGGAAAGTTCTCCT 1139
QY 1262 CCAAAACAATGAAGAAGTTTTCAGGCACCTGCAAGAGAGACCTGCGTGGAAATGTCAGAG 1321
Db 1140 TCCAAACACAGGAAGAAGTTTTCAGGCCCGGCAAGAGAGAGCTGCGTGGAGTGTGAGAAG 1199
QY 1322 ACAGTCTATCAATGGAGCGCTCTTTGGCCAAACAGCAGAGGTGTTTCAACATCAGCTGCTTC 1381
Db 1200 ACGGTGTACCCCATGGAAACGGCTCTTGGCCAAACAGCAGAGGTGTTTCAACATCAGCTGTTTC 1259
QY 1382 CGTTGCTCTATTGGCAACAACTCAGTCTAGGAACATATGCAATCTTTACATGGAAGA 1441
Db 1260 CGATGCTCTCTACTGCAACAACTCAGTCTAGGAACATATGCAATCTTTACATGGAAGA 1319
QY 1442 ATCTATTGTAAGCTCACTTCAATCACTCTTTTAAATCTTAAGGGCACTATGATGAAGCC 1501
Db 1320 ATCTACTGTAAGCTCACTTCAATCACTCTTTCAAGTCCAAAGGCACTATGACGAGAGC 1379
QY 1502 TTTGGGCACAGACACACAAAGGATCTATGGGCAAGCAAAATGAAACGAAGAGATTTTG 1561
Db 1380 TTTGGGCACAGCAGCATAAGATCTGTGGGCAAGCAAGAGCGACAAATGAGGAGACTTTTG 1439
QY 1562 GAGAGACCCAGCTGTGCAAAATGCAAGGGAGAGCCCTCACAGCCAGGGGTAGAAGAT 1621
Db 1440 GGGAGACCCAGCCAGCCGCTAATGAGGGAGAGCCCTCATAGCCCGGTGTAGAGAT 1499
QY 1622 GCCCTATTGCTAAGTGGGTGCTTCCGGTGCAGATGATGAAGCCCAAGGCCCTCTCTCTAG 1681
Db 1500 GCCCCATGCCCAAGTGGCGTGTGGGCAAGTATGGAAGCCCAAGGCCCTCTCTCTAG 1559
```

```
QY 1682 CAGGAAAGGAAGACAAAGCCAGCTGAACCAAGAAAGCTGAGGATGCCCTGGCCACCCCCC 1741
Db 1560 AGGAGAGGGAGGATGAAGCTCTCGGAGACCAAGAAAGCTGAGGATGCCCTGGCTCCGCCA 1619
QY 1742 ACTGAACCTTGGAAAGTTTCAGAAAGTCCCTTGGAGGAGGATCAAAATGTCAAAAGCCCAA 1801
Db 1620 GCCGAGCTGGCGGTTTCCGAAAGTCCCTTGGAGGAGGATCAAAATGTCAAAAGCCCAA 1679
QY 1802 TGGCTCTCTGAAGACGAAAATCAGCAAGCCCGCAAGTTCTCTGAGGATGTCGATCTAGATCTG 1861
Db 1680 TGGCTCTCGGAGGATGACGCTCTGCAAGCGGAGGCCCGGAGGATGTAGATCTGACCTG 1739
QY 1862 AAGAAGCTTAAGACGATCTTCTTCACTGAAGGAAAGAACGCCGCCATTTCACTGTAGCAGCT 1921
Db 1740 AAGAAGCTTCGGCGGTTTCTGCTCCCTGAAAGGAGAGAGCCGCCCTTTACGCTAGCAGCT 1799
QY 1922 TCATTTCAAAGCACCTCTGTCGAAGGCCCAAAATCTGTGTCCCACTATACGAAGAAAGCC 1981
Db 1800 TCTTTTCGAACCTCTTCCATCAAGAGCCCAAGGCCCTCTTCGCCATCTCTCAGGAAAGGT 1859
QY 1982 TGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGTGGAAAGAGTTGCA---GAAAGGAAA 2038
Db 1860 TGGAGCGAGTCCGAGCAGAGTGAAGAGTTTGGAGGAGGATAGCAGCAGATGGAAGGAAA 1919
QY 2039 CAAGTGGAAATGCCAAGGCTTCTTAAGAAAGATGGGAATGTGGGAAACAAACCTGGCAA 2098
Db 1920 CAAACGGAAATGCCAAGGCTCTCGGGAGAAAGAGATGTGGGAAATCAACCTGGCAG 1979
QY 2099 AACAAGAAATCTAAAGGAGAGACAGGAGAGAGAAAGTAAGGAAGGTCTAGTTTGGAGATG 2158
Db 1980 -----GGCGAGGAAGTGGCGAGGAGCAAGGACCGTAGTTCTTTTTCAGCTG 2024
QY 2159 GAGATGAGAAATCTTTAGAAAATGGTGCAGACTCCGATGAAGATGATAAGCTTCTCTC 2218
Db 2025 GAGGTGAGAAATTTTATGGAAAATGGAGCAACATAGCTGAAGATGACAACCATGTCCAC 2084
QY 2219 AAACAACAAATCTCCACAAGAACCCCAAGTCTCTGAATTTGGTTCGAGTTTGTAGACAACAC 2278
Db 2085 GCACAGCAGTCCCGTTAGAACCCGAGGCTCGGTTTGGTCTGGCTTTGTAGACACCAG 2144
QY 2279 TTTGTGAAGAAATCACTACTCAGAAATCAGAAATCCCAAGGATGTGGAATCTTGGAGGGA 2338
Db 2145 GCTGTAAAGAAATCACTACCCAGAAATCAAAATCCCAAGATGTCCGGTTCTGGAGGGA 2204
QY 2339 GAAAGTGTCAAGAGCTCTCTGTGGAAGAACAGATAAGAGAAATCGGTATTATGATGAG 2398
Db 2205 GAAAGTGTCCGAGAGCTCTCTGTGGAAGAACAGATAAGAGAAACCCGCTACTATCATGAG 2264
QY 2399 GATGAGGATGAAGAGTGAACAAATGCAATGATGTGGGCCCTTAAATTCATGTTAGTTTA 2458
Db 2265 GACGAGGATGAAGAAATGACCCGCGCCATGGTTCTGGCC-----TTAGATTTTGGTTA 2317
QY 2459 GCGAGCCACTGCGCTTTTGTCAAAATGTGATGCACATAAGCAGGTATCCAGCATGAAATG 2518
Db 2318 GCACTCTACTGCCCTTTATCAA---GTGGTGGCAGAAACAGGAGCTTGGCGCCAGGG 2374
QY 2519 TAATTTACTTGGAAAGTAACTTTGGAAAGAAATTCCTTCT-TAAAAATCAAAAACAAAACAA 2577
Db 2375 TCATTTCATGTGGAGCACTTTTGAATGAATTCCTGCTAGCAAAACAGGCGCAAAAACAA 2434
QY 2578 AAAACACAAAAACACATTTCTAAAT 2603
Db 2435 GCAAAAACAAAAACAGACAAACAACT 2460
```

## RESULT 5

AL578990/c

LOCUS

DEFINITION

AL578990 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DK012VC22 3-PRIME, mRNA sequence.

ACCESSION

AL578990

VERSION

AL578990.2 GI:31317177

KEYWORDS

EST.

AL578990 1014 bp mRNA linear EST 01-JUN-2003

AL578990 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens

cDNA clone CS0DK012VC22 3-PRIME, mRNA sequence.

AL578990

AL578990.2 GI:31317177

EST.





Db 181 GGTGATATCCCCACATCAAGACCGTGAGGATCTTAAAGCCACTCAACAGAAAGTAAA 240  
Qy 614 AAAATGGAAAATTTGTCTAGGAAATCCAGGATGAAGTAGAAAAATCAGAAATCAGTCAA 673  
Db 241 AAAATGGAAAATTTGTCTAGGAAATCCAGGATGAAGTAGAAAAATCAGAAATCAGTCAA 300  
Qy 674 AACACAGATGCTTCGGGCAAAATAGAGAAATATAATGTTCGCTGAAACAGGCTTAAGATG 733  
Db 301 AACACAGATGCTTCGGGCAAAATAGAGAAATATAATGTTCGCTGAAACAGGCTTAAGATG 360  
Qy 734 ATGTTTGACAAAGGTGACCACTCAAACTAGATTTCTCCGGGCCCAAGCGGAAGTGCA 793  
Db 361 ATGTTTGACAAAGGTGACCACTCAAACTAGATTTCTCCGGGCCCAAGCGGAAGTGCA 420  
Qy 794 AGTGAAGCAAGATCTCTGMAAACAGCTATTCTCTAGATGACCTGGAATATAGGCCAGGT 853  
Db 421 AGTGAAGCAAGATCTCTGMAAACAGCTATTCTCTAGATGACCTGGAATATAGGCCAGGT 480  
Qy 854 CAGTTGTATCTTCTACATTTGACTCGGAGAAAAATAGAGATAGCAAAATCTGGAACCT 913  
Db 481 CAGTTGTATCTTCTACATTTGACTCGGAGAAAAATAGAGATAGCAAAATCTGGAACCT 540  
Qy 914 CCACGCTCTCAGAACCTCTATAAGGATCGAATGGCCAAAGTACCAGGACGCTGTGTCC 973  
Db 541 CCACGCTCTCAGAACCTCTATAAGGATCGAATGGCCAAAGTACCAGGACGCTGTGTCC 600  
Qy 974 AAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGTGGCGAAATCAAA 1033  
Db 601 AAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGTGGCGAAATCAAA 660  
Qy 1034 ATTCATAAATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTCTGAGTCTGCATCACCCAT 1093  
Db 661 ATTCATAAATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTCTGAGTCTGCATCACCCAT 720  
Qy 1094 CAGGAAGGGGAAAAGATTCTGCAAAATAGAAATAGCCCTGGCAGTCCGTTCAACCCCTGCC 1153  
Db 721 CAGGAAGGGGAAAAGATTCTGCAAAATAGAAATAGCCCTGGCAGTCCGTTCAACCCCTGCC 780  
Qy 1154 GAAGATGACTCCCGTACTCCAGGTTAAGAGTGAAGTTCAACAGCTGTCCATCCCAG 1213  
Db 781 GAAGATGACTCCCGTACTCCAGGTTAAGAGTGAAGTTCAACAGCTGTCCATCCCAG 840  
Qy 1214 CCACAAAGTTCAGGACCTCCAGACCTCCAG-TCTTTCTGAAAGTTCTCTCCCAAGCAAT 1272  
Db 841 CCACAAAGTTCAGGACCTCCAGACCTCCAG-TCTTTCTGAAAGTTCTCTCCCAAG-AAT 899  
Qy 1273 GAAGAAATTTTCAGGACCTCCAGACCTCCAGACCTCCGTTGAA--TGTGAGAGACAGTCTA 1329  
Db 900 GGAAGATTTTCAGGACCTCCAGACCTCCAGACCTCCGTTGAA--TGTGAGAGACAGTCTA 959  
Qy 1330 TCCATGG-AGCGTCTCTTGGCCACACAGGAGTG 1363  
Db 960 TCCCATGGAAGCGTCTCTTGGCCAAACAGGAGGG 994

RESULT 8  
BI871114  
LOCUS  
DEFINITION BI871114 880 bp mRNA linear EST 11-OCT-2001  
mRNA sequence.  
ACCESSION BI871114  
VERSION BI871114.1 GI:16044787  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 880)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1AM12030 row: f column: 23  
High quality sequence stop: 799.  
Location/Qualifiers  
1. 880  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5404054"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 90"  
/notes="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH MGC Library."  
BASE COUNT 308 a 136 c 181 g 255 t  
ORIGIN  
Query Match 21.8%; Score 806.8; DB 12; Length 880;  
Best local Similarity 96.7%; Pred. No. 3e-94;  
Matches 845; Conservative 0; Mismatches 27; Indels 2; Gaps 2;  
Qy 2124 GGAAGAGAGTAGAGGAGGTCTAGTTTGGAGATGAGAGATGAGAAATG 2183  
Db 1 GGAAGAGAGTAGAGGAGGTCTAGTTTGGAGATGAGAGATGAGAAATG 60  
Qy 2184 GTGCAGACTCCGATGAAGATGATAACAGCTTCTCTCAACAAACAACTCTCCACAGAAACCCA 2243  
Db 61 GTGCAGACTCCGATGAAGATGATAACAGCTTCTCTCAACAAACAACTCTCCACAGAAACCCA 120  
Qy 2244 AGTCTCTGAATTTGGTTCGAGTTTGTAGACAAACACCTTCTCTGAAGAAATTCACCTACTCAGA 2303  
Db 121 AGTCTCTGAATTTGGTTCGAGTTTGTAGACAAACACCTTCTCTGAAGAAATTCACCTACTCAGA 180  
Qy 2304 ATCAGAAATCCAGGATGTGGAACTCTGGAGGGAGAAAGTGTCAAGAGCTCTCTGTGG 2363  
Db 181 ATCAGAAATCCAGGATGTGGAACTCTGGAGGGAGAAAGTGTCAAGAGCTCTCTGTGG 240  
Qy 2364 AAGAACAGATAAAGAGAAATCCGTTATTATGATGAGGATGAGGATGAGAGTACAAATG 2423  
Db 241 AAGAACAGATAAAGAGAAATCCGTTATTATGATGAGGATGAGGATGAGAGTACAAATG 300  
Qy 2424 CAATGATGCTGGGCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAAAT 2483  
Db 301 CAATGATGCTGGGCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAAAT 360  
Qy 2484 GTGATGCACATAGCAGGTATCCAGCATGAAATGTAATTTACTTTGGAAGTAACTTTGGA 2543  
Db 361 GTGATGCACATAGCAGGTATCCAGCATGAAATGTAATTTACTTTGGAAGTAACTTTGGA 420  
Qy 2544 AAGAAATTCCTTTTAAATCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2603  
Db 421 AAGAAATTCCTTTTAAATCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 480  
Qy 2604 ACTAGAGATAACTTTACTTTAAATTTCTTTTACTGATGATGATGATGATGATGATGATGATGAT 2663  
Db 481 ACTAGAGATAACTTTACTTTAAATTTCTTTTACTGATGATGATGATGATGATGATGATGATGAT 540  
Qy 2664 AAGGCTTGTAAGTGGGAAATATCCACCTGATAATAGCCAGATTTCTACTGATTTCCCA 2723  
Db 541 AAGGCTTGTAAGTGGGAAATATCCACCTGATAATAGCCAGATTTCTACTGATTTCCCA 600  
Qy 2724 AAGGCAATTTAAGGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2783



Db 601 AAAGCAATATTAAGGTAGATGATTAGTAGTATATATGTTACACACTATTTTGAAT 660

QY 2784 TAGAGACATACAGAGGAATTTAGGGCTTTAAACA-TTACAGCTGAATGCACCTTTAGTA 2842

Db 661 TAGAGACATACAGAGGAATTTAGGGCTTTAAACA-TTACAGCTGAATGCACCTTTAGTA 720

QY 2843 TAAAGGACACAGTTTGTATATTTTAAATGAATACCAATTAA-TTTTGTAGTATTTACC 2901

Db 721 TAAAGGACACAGTTTGTATATTTTAAATGAATACCAATTAA-TTTTGTAGTATTTACC 780

QY 2902 TGTTAAGAGATTTATTTAGTCTTTTAAATTTTAAATTTTAAATTTTCTGCTGTGATATAT 2961

Db 781 TGTTAAGAGATTTATTTAGTCTTTTAAATTTTAAATTTTAAATTTTCTGCTGTGATATAT 840

QY 2962 GAGGAATTTACTACTTTATGTCCTGCTCTCTAAA 2995

Db 841 GCAGCAATTAAGTCTTATGTCCTGCTCTCTAAA 874

RESULT 9

CD171644

LOCUS

DEFINITION AGENCOURT 14065890 NIH\_MGC\_180 Homo sapiens cDNA clone

IMAGE:30378885 5', mRNA sequence.

ACCESSION CD171644

VERSION CD171644.1 GI:30852393

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 884)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM445 row: g column: 22  
High quality sequence start: 16  
High quality sequence stop: 648.  
Location/Qualifiers  
1. 884  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30378885"  
/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_180"  
/notes="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning ). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC Library."  
306 a 168 c 225 g 185 t

BASE COUNT 306 a 168 c 225 g 185 t

ORIGIN

Query Match 21.7%; Score 802.4; DB 14; Length 884;  
Best Local Similarity 98.4%; Pred. No. 1.1e-93;  
Matches 842; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 1730 TGCCACCCCCCACTGAATTTGGAAGTTCAGGAAGTGCCTTGAGCAAGGGATCAAAATG 1789

Db 27 TGCCACCCCCCACTGAATTTGGAAGTTCAGGAAGTGCCTTGAGCAAGGGATCAAAATG 86

QY 1790 TCAAGCCCAAAATGGCTCTCTGAGACGAAATCAGCAAGCCCGAAGTTCCTCAGGATGTC 1849

Db 87 TCAAGCCCAAAATGGCTCTCTGAGACGAAATCAGCAAGCCCGAAGTTCCTCAGGATGTC 146

QY 1850 GATCTAGATCTGAAGAAGCTTAAGAGCATCTTCTTCACTGAAGAAAGAGCGCCCATTC 1909

Db 147 GATCTAGATCTGAAGAAGCTTAAGAGCATCTTCTTCACTGAAGAAAGAGCGCCCATTC 206

QY 1910 ACTGTAGCAGCTTCATTTTCAAGACCTCTGTCAAGAGCCCAAAAACCTGTGTCCTCCACCT 1969

Db 207 ACTGTAGCAGCTTCATTTTCAAGACCTCTGTCAAGAGCCCAAAAACCTGTGTCCTCCACCT 266

QY 1970 ATCAGAAAGGCTGGAGCATGTCTAGAGCAGATGAAGAGTCTGTGGTGGAGAGTGTGA 2029

Db 267 ATCAGAAAGGCTGGAGCATGTCTAGAGCAGATGAAGAGTCTGTGGTGGAGAGTGTGA 326

QY 2030 GAAAGGAAACAAGTGCAGAAATCCAGAGGCTTCTAAGACAAATGGGAATGTGGGAAAAACA 2089

Db 327 GAAAGGAAACAAGTGCAGAAATCCAGAGGCTTCTAAGACAAATGGGAATGTGGGAAAAACA 386

QY 2090 ACCTGGCAAAAACAAGAAATCTTAAAGGAGAGACAGAGGAAAGTAAGGAAGTCTATAGT 2149

Db 387 ACCTGGCAAAAACAAGAAATCTTAAAGGAGAGACAGAGGAAAGTAAGGAAGTCTATAGT 446

QY 2150 TTGGAGATGGAGAAATGAGAAATCTTGTAGAAAAATGGTGCAGACTCCGATGAAGATGATAAC 2209

Db 447 TTGGAGATGGAGAAATGAGAAATCTTGTAGAAAAATGGTGCAGACTCCGATGAAGATGATAAC 506

QY 2210 AGCTTCTCTCAAAACAACAATCTCCACAAGAACCCAAAGTCTCTGAATTTGGTCGAGTTTGT 2269

Db 507 AGCTTCTCTCAAAACAACAATCTCCACAAGAACCCAAAGTCTCTGAATTTGGTCGAGTTTGT 566

QY 2270 GACAACACCTTTGCTGAAGAAATCTACTCTCAGAAATCAGAAATCCAGATGTGGAACCT 2329

Db 567 GACAACACCTTTGCTGAAGAAATCTACTCTCAGAAATCAGAAATCCAGATGTGGAACCT 626

QY 2330 TGGAGGGAGAAAGTGGTCAAAAGAGCTCTCTGTGGAAGAAACAGATAAAGAGAAATCGGTAT 2389

Db 627 TGGAGGGAGAAAGTGGTCAAAAGAGCTCTCTGTGGAAGAAACAGATAAAGAGAAATCGGTAT 686

QY 2390 TATGATGAGGATGAGGATGAAGAGTGAACAAATGCAATGATCTGGGCTTTAAATTCATG 2449

Db 687 TATGATGAGGATGAGGATGAAGAGTGAACAAATGCAATGATCTGGGCTTTAAATTCATG 746

QY 2450 TTAGTGTAGGAGGACCTGCCCC-TTGTCAAAATGTGATGCACATAAGCAGGTATCCCA 2508

Db 747 TTAGTGTAGGAGGACCTGCCCC-TTGTCAAAATGTGATGCACATAAGCAGGTATCCCA 806

QY 2509 GCATCAAAAT-GTAATTTTACTTTGGAAGTAACTTTTGGAAAGTAACTTTGGGAAAGAAATTCCTTCTTAAATCAA 2566

Db 807 GCATCAAAATGGAATTTTACTTTGGAAGTAACTTTTGGGAAAGAAATTCCTTCTTAAATTCG 866

QY 2567 AAACAAAAACAAAAA 2582

Db 867 AAACAAAAACAAAAA 882

RESULT 10

BM914155

LOCUS

DEFINITION AGENCOURT\_6608184 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5479040

5', mRNA sequence.

ACCESSION BM914155

VERSION BM914155.1 GI:19364534

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1142)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLC1999 row: k column: 09  
High quality sequence stop: 878.  
Location/Qualifiers  
1. 1142  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5479040"  
/tissue\_type="astrocytoma grade IV, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC\_98"  
/note="Organ: brain; Vector: pOTB7; Site: 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH MGC Library."

BASE COUNT 376 a 282 c 276 g 208 t

ORIGIN

Query Match 21.3%; Score 787.8; DB 12; Length 1142;  
Best Local Similarity 94.9%; Pred. No. 7.4e-92;  
Matches 890; Conservative 0; Mismatches 37; Indels 10; Gaps 6;

QY	347	GTGTTAAAGAGAACTGGGAGAACCCAGGCTGGGAGCAGAGTCTCACAGACTCTCTA	406
Db	1	GTGTTAAAGAGAACTGGGAGAACCCAGGCTGGGAGCAGAGTCTCACAGACTCTCTA	60
QY	407	CGGAACAGCAGCACTAGATTAGGCACAGACAGACCATCTCTGCTGAAGTCACAAAGC	466
Db	61	CGGAACAGCAGCACTAGATTAGGCACAGACAGACCATCTCTGCTGAAGTCACAAAGC	120
QY	467	CACGCTGCTTCTGAGCCAAAGCTGACCAAGAAACAAATCCACCCAGACTTAGACTC	526
Db	121	CACGCTGCTTCTGAGCCAAAGCTGACCAAGAAACAAATCCACCCAGACTTAGACTC	180
QY	527	AGGTCACCTCTGAGGCCCTCTGTCAGGTCGATATCCCAATCAAGGCGGTGAGGAT	586
Db	181	AGGTCACCTCTGAGGCCCTCTGTCAGGTCGATATCCCAATCAAGGCGGTGAGGAT	240
QY	587	CTTAAAGACCACTCAACAGAAAGTAAAAAATGGAATAATGCTAGGAGATCCAGGCAT	646
Db	241	CTTAAAGACCACTCAACAGAAAGTAAAAAATGGAATAATGCTAGGAGATCCAGGCAT	300
QY	647	GAAGTAGAAAAATCAGAAATCAGTGAACACAGATGCTTCGGGCAAAATAGAGAAATAT	706
Db	301	GAAGTAGAAAAATCAGAAATCAGTGAACACAGATGCTTCGGGCAAAATAGAGAAATAT	360
QY	707	AATGTTCCGCTCAGACAGCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAG	766
Db	361	AATGTTCCGCTCAGACAGCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAG	420
QY	767	ATTCTCCGGGCCAAAGCCGAGTGAAGTGAAGGAGATCTCTGAAGAACAGCTATTCT	826
Db	421	ATTCTCCGGGCCAAAGCCGAGTGAAGTGAAGGAGATCTCTGAAGAACAGCTATTCT	480
QY	827	CTAGATGACCTGGAAATAGGCCAGGTCAGTTGTTCATCTTCTACATTGACTCGGAAA	886
Db	481	CTAGATGACCTGGAAATAGGCCAGGTCAGTTGTTCATCTTCTACATTGACTCGGAAA	540
QY	887	AATGAGATACGAAATCTGGAACTTCCAGCTCTCAGAAACCTCTATAAAGATCGA	946
Db	541	AATGAGATACGAAATCTGGAACTTCCAGCTCTCAGAAACCTCTATAAAGATCGA	600

QY	947	ATGCCCAAGTACCAGCAGCTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAATGAG	1006
Db	601	ATGCCCAAGTACCAGCAGCTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAATGAG	660
QY	1007	CTCAAGCCAGTGGTGGCAAAATCAAAATTCATAAAATGGACAAAGAGAGATGTGCC	1066
Db	661	CTCAAGCCAGTGGTGGCAAAATCAAAATTCATAAAATGGACAAAGAGAGATGTGCC	720
QY	1067	CCAGGT-CTGAGGTCTGCATCACCCTCAGCAAGGGGAAAAGATTCTGCAAAATGAGAA	1125
Db	721	CCAGGTCTCCCTGAGGTCTGCATCACCCTCAGCAAGGGGAAAAGATTCTGCAAAATGAGAA	780
QY	1126	TAG-CTGCGCAGTCCGTTCCACCCC--TCCGGAAGATGACTCCCGTGACTCCCAAGTT--	1180
Db	781	TAGCCCTGCGCAGTCCGTTTCCACCCCCTGCGGAATAATGACTCCCGGACTCCCAAGTTA	840
QY	1181	-AAGATGAGGTTCAACAGCCTGTCCATCCCAA-GCCATAGTGCAGATT--CCAGAGC	1236
Db	841	AAAGGGAGGTTTCAACAGCCTGGCCATCCCAAGGCCCTTAAGTCCAGATTTCCAAAGGC	900
QY	1237	CTCCAGTCTTCTGAAAGTTTCTCCTCC	1263
Db	901	CTCCAGGCTTTTGAAGGTTCTCC	927

RESULT 11  
BM721370

LOCUS BM721370 810 bp mRNA linear EST 01-MAR-2002  
DEFINITION UI-E-E01-aib-f-18-0-UI.r1 UI-E-E01 Homo sapiens cDNA clone  
UI-E-E01-aib-f-18-0-UI 5', mRNA sequence.

ACCESSION BM721370

VERSION BM721370.1

KEYWORDS GI:19041235

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 810)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

PUBMED 8889548

COMMENT

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1. 810

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-E-E01-aib-f-18-0-UI"

/tissue\_type="fetal eye"

/dev\_stage="fetal"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-E01 is a normalized cDNA library containing the

following tissue(s): fetal eye. The library was

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CGCGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."										
BASE COUNT		277 a	124 c	155 g	251 t	3 others				
ORIGIN										
Query Match		20.9%; Score 775.4; DB 12; Length 810;								
Best Local Similarity		98.6%; Pred. No. 3.3e-90;								
Matches 801; Conservative		0;	Mismatches	9;	Indels	2;	Gaps	2;		
QY	2308	GAATATCCAGGATGTGAACTCTCTGGGAGGAGAAGTGGTCAAAGAGCTCTCTGTGGAAGA	2367							
Db	1	GAATATCCAGGATGTGAACTCTCTGGGAGGAGAA-TGGTCAAAGAGCTCTCTGTGGAAGA	59							
QY	2368	ACAGATAAAGAGAATCGGTATTATGATGAGGATGAGGATCGAGATGCAAGATGCAAAATTCGAAT	2427							
Db	60	ACAGATAAAGAGAATCGGTATTATGATGAGGATGAGGATGCAAGATGCAAAATTCGAAT	119							
QY	2428	GATGCTGGGCCTTAAATTCATGTTAGTGTGTAGCGAGCCACTGCGCCTTTGTCAAAATGTGA	2487							
Db	120	GATGCTGGGCCTTAAATTCATGTTAGTGTGTAGCGAGCCACTGCGCCTTTGTCAAAATGTGA	179							
QY	2488	TGCACATAAGCAGGTATCCGAGCATGAAATGTAAATTTACTTGGAAATGACCTTTGGAAAAG	2547							
Db	180	TGCACATAAGCAGGTATCCGAGCATGAAATGTAAATTTACTTGGAAATGACCTTTGGAAAAG	239							
QY	2548	AATTCCCTCTTAAATCAAAAACAAAACAAAACAAAACAAAACAAAACACATTTCTAAATACTA	2607							
Db	240	AATTCCCTCTTAAATCAAAAACAAAACAAAACAAAACAAAACAAAACACATTTCTAAATACTA	299							
QY	2608	GAGATAACTTTACTTAAATTTCTTCATTTTAGCAGTGTATGATATGCATAAGTGTCTGTAAGG	2667							
Db	300	GAGATAACTTTACTTAAATTTCTTCATTTTAGCAGTGTATGATATGCATAAGTGTCTGTAAGG	359							
QY	2668	CTTGTAACCTGGGAAATATTCCACCTGATTAATAGCCGAGATTTCTACTGTATTTCCCAAAAG	2727							
Db	360	CTTGTAACCTGGGAAATATTCCACCTGATTAATAGCCGAGATTTCTACTGTATTTCCCAAAAG	419							
QY	2728	GCAATATTAGGTAGATAGATGATTAGTAGTATATGTTTACACACTATTTTGGAAATTAGA	2787							
Db	420	GCAATATTAGGTAGACAGATGATTAGTAGTATATGTTTACACACTATTTTGGAAATTAGA	479							
QY	2788	GAACATACAGAAGAAATTTAGGGGCTTTAAACATTTACGACTGAATGACATTGATATAAAG	2847							
Db	480	GAACATACAGAAGAAATTTAGGGGCTTTAAACATTTACGACTGAATGACATTGATATAAAG	539							
QY	2848	GGCACAGTTTGATATTTTAAATGAATACCAATTTTAAATTTTAGTATTTTACTCTGTTAA	2907							
Db	540	GGCACAGTTTGATATTTTAAATGAATACCAATTTTAAATTTTAGTATTTTACTCTGTTAA	599							
QY	2908	GAGATATTAGTCTTTAAATTTTTTAGGTTAAATTTTCTTGCTGTGATATATATGAGAA	2967							
Db	600	GAGATATTAGTCTTTAAATTTTTTAGGTTAAATTTTCTTGCTGTGATATATATGAGAA	659							
QY	2968	TTTACTACTTTATGTCTGCTCTTAACTACATCTGTAACCTGAGCTCGAGTCTGAGGTATAA	3027							
Db	660	TTTACTACTTTATGTCTGCTCTTAACTACATCTGTAACCTGAGCTCGAGTCTGAGGTATAA	719							
QY	3028	TACAACAGAGCACTTTTGTAGGCAATGAAAAACCAACCTACACTCTTCGGTGCTTTAGAG	3087							
Db	720	TACAACAGAGCACTTTTGTAGGCAATGAAAAACCA-CTACACTCTTCGGTGCTTTAGAG	778							
QY	3088	AGATCTGCTGTCTCCCAATAAAGCTTTTGAT	3119							
Db	779	AGATCTGCTGTCTCCCAATAAAGCTTTTGAT	810							

RESULT 12	602633701F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4778761 5'	785 bp	mRNA	linear	EST 15-MAY-2001
LOCUS	602633701F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4778761 5'				
DEFINITION	mRNA sequence.				
ACCESSION	BG740815				
VERSION	BG740815.1	GI:114051468			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 785)				
JOURNAL	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgabbs-r@mail.nih.gov">cgabbs-r@mail.nih.gov</a> Tissue Procurement: James Cleaver, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM10635 row: a column: 02 High quality sequence stop: 785. Location/Qualifiers 1..785 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:4778761" /lab_host="DH10B (Tl phage-resistant)" /clone_lib="NCI_CGAP_Skn3" /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: Not1; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library." BASE COUNT 240 a 137 c 134 g 274 t ORIGIN Query Match 20.9%; Score 774; DB 10; Length 785; Best Local Similarity 99.9%; Pred No. 5e-90; Matches 785; Conservative 0; Mismatches 0; Indels 1; Gaps 1; QY 2607 AGAGATAACTTTACTTAAATCTCTTATTTAGCAGTGATGATATGCATAGTCTGTAAAG 2666 Db 1 AGAGATAACTTTACTTAAATCTCTTATTTAGCAGTGATGATATGCATAGTCTGTAAAG 60 QY 2667 GCTTGTAACTGGGAAATATTCACCTGTAATATAGCCAGATTTCTGTGTTATTCACAAA 2726 Db 61 GCTTGTAACTGGGAAATATTCACCTGTAATATAGCCAGATTTCTGTGTTATTCACAAA 120 QY 2727 GGCAATATTAAGTAGATAGATTAGTAGTATTTGTTACACACTATTTTGGAAATAG 2786 Db 121 GGCAATATTAAGTAGATAGATTAGTAGTATTTGTTACACACTATTTTGGAAATAG 180 QY 2787 AGAACATACAGAAGGAATTTAGGGCTTTAAACATTACGACTGAATGCACTTTAGTATAAA 2846 Db 181 AGAACATACAGAAGGAATTTAGGGCTTTAAACATTACGACTGAATGCACTTTAGTATAAA 240 QY 2847 GGCACAGTTTGTATATTTTAAATGAATACCAATTTAAATTTTGTAGTATTACCTGTTA 2906 Db 241 GGCACAGTTTGTATATTTTAAATGAATACCAATTTAAATTTTGTAGTATTACCTGTTA 300 QY 2907 AGAGATTATTTAGTCTTTAAATTTTTTGGTTAAATTTCTTCTGTGATATATAGGA 2966 Db 301 AGAGATTATTTAGTCTTTAAATTTTTTGGTTAAATTTCTTCTGTGATATATAGGA 360 QY 2967 ATTTTACTATTTATGCTGCTCTCTAAACTACATCTCCGAACTCGACGCTCTGAGGTATA 3026				

||||| 361 ATTACTACTTTATGTCCTGCTCTTAATACTACCTGAACTGACGCTCGAGGTATA 420  
||||| QY 3027 ATACAAACAGAGCAGCTTTTGGAGCAATGGAAGAAACCACTACACTCTTCGGTGCTTAGA 3086  
||||| Db 421 ATACACACAGCAGCTTTTGGAGCAATGGAAGAAACCACTACACTCTTCGGTGCTTAGA 480  
||||| QY 3087 GAGATCTGCTGCTCCCAATAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAA 3146  
||||| Db 481 GAGATCTGCTGCTCCCAATAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAA 540  
||||| QY 3147 TGATGCTTTCTTTCTGCTGATATCTGCTCTCATTAATTTACTGAAAGCTGCAATATT 3206  
||||| Db 541 TGATGCTTTCTTTCTGCTGATATCTGCTCTCATTAATTTACTGAAAGCTGCAATATT 600  
||||| QY 3207 TTAGTAATACCTTCGGATCACTGCTCCCATCTTCCTGCTTTAGAGCAAGTGAAGAGTT 3266  
||||| Db 601 TTAGTAATACCTTCGGATCACTGCTCCCATCTTCCTGCTTTAGAGCAAGTGAAGAGTT 660  
||||| QY 3267 TAAAGGAGGAAGAAAGAACTGCTTTACACACTTGAGCTTCAGACCTCTAAACCCTGT 3326  
||||| Db 661 TAAAGGAGGAAGAAAGAACTGCTTTACACACTTGAGCTTCAGACCTCT-AAACCCTGT 719  
||||| QY 3327 ATTTCCTTATGATGCTCCCTTTTGGACACACTAATTTTAAATACTTACTAGCTCTGAA 3386  
||||| Db 720 ATTTCCTTATGATGCTCCCTTTTGGACACACTAATTTTAAATACTTACTAGCTCTGAA 779  
||||| QY 3387 ATATAT 3392  
||||| Db 780 ATATAT 785

RESULT 13  
AQ314676/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AQ314676 777 bp DNA linear GSS 04-MAY-1999  
RPC111-103F24-TV RPCI-11 Homo sapiens genomic clone RPCI-11-103F24,  
genomic survey sequence.  
AQ314676  
AQ314676.1 GI:4046139  
GSS.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.  
Use of human BAC End Sequences for Sequence-Ready Map Building  
Unpublished  
Other GSSs: RPC111-103F24.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genetics ([http://www.tigr.org/tldb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html))  
Seq primer: T

Class: BAC ends.  
Location/Qualifiers  
1. .777  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:753311"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-103F24"  
/sex="Male"

FEATURES  
source

/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPC111 Human Male BAC Library"  
BASE COUNT 164 a 199 c 139 g 275 t  
ORIGIN  
Query Match 20.7%; Score 766.4; DB 28; Length 777;  
Best Local Similarity 99.2%; Pred. No. 4.8e-89;  
Matches 770; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1775 GAAGGGATCAAAATGTCAAAGCCCAATGCGCTCTGAAAGCAAAATCAGCAAGCCCGAA 1834  
Db 777 GAAGGGATCAAAATGTCAAAGCCCAATGCGCTCTGAAAGCAAAATCAGCAAGCCCGAA 718  
QY 1835 GTTCCTCAGATGTCGATCTAGATCTCAAGAACTAAGACGATCTTCTTCACTGAAGGAA 1894  
Db 717 GTTCCTCAGATGTCGATCTAGATCTCAAGAACTAAGACGATCTTCTTCACTGAAGGAA 658  
QY 1895 AGAAGCGCCCATTCACCTGTAGCAGCTTCAATTTCAAAGCACCTCTGTCAAGAGCCCAAAA 1954  
Db 657 AGAAGCGCCCATTCACCTGTAGCAGCTTCAATTTCAAAGCACCTCTGTCAAGAGCCCAAAA 598  
QY 1955 ACTGTGTCCCACTATCAGGAAAGGCTGGAGCATGTTCAGAGCAGAGTGAAGAGTCTGTG 2014  
Db 597 ACTGTGTCCCACTATCAGGAAAGGCTGGAGCATGTTCAGAGCAGAGTGAAGAGTCTGTG 538  
QY 2015 GGTGGAAGATTCAGAAAGCAAACTGAAGTGGAAATGCCAAGGCTTCTTAAGAGAAATGGG 2074  
Db 537 GGTGGAAGATTCAGAAAGCAAACTGAAGTGGAAATGCCAAGGCTTCTTAAGAGAAATGGG 478  
QY 2075 AATGTGGGAAAAACAACCTGCGAAAAACAAGATCTTAAGAGCAGACAGGAGAGAGT 2134  
Db 477 AATGTGGGAAAAACAACCTGCGAAAAACAAGATCTTAAGAGCAGACAGGAGAGAGT 418  
QY 2135 AAGGAAGCTCATAGTTTGGAGATGGAGAAATGAGAATCTTTAGAAAAATGGTCAGACTCC 2194  
Db 417 AAGGAAGCTCATAGTTTGGAGATGGAGAAATGAGAATCTTTAGAAAAATGGTCAGACTCC 358  
QY 2195 GATGAAGATGATACAGCTTCTCAAAACAACAACTCTCCACAAGACCCCAAGTCTCTGAAT 2254  
Db 357 GATGAAGATGATACAGCTTCTCAAAACAACAACTCTCCACAAGACCCCAAGTCTCTGAAT 298  
QY 2255 TGCTCGAGTTTCTAGACACACCTTTGCTGAAGAATTCACACTACTCAGAAATCAGAAATCC 2314  
Db 297 TGCTCGAGTTTCTAGACACACCTTTGCTGAAGAATTCACACTACTCAGAAATCAGAAATCC 238  
QY 2315 CAGGATGTGAACTCTGGGAGGAGAGTGGTCAAAAGAGCTCTCTGTGGAAGAACAGATA 2374  
Db 237 CAGGATGTGAACTCTGGGAGGAGAGTGGTCAAAAGAGCTCTCTGTGGAAGAACAGATA 178  
QY 2375 AAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGAACAAATTCGAATGATGCTG 2434  
Db 177 AAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGAACAAATTCGAATGATGCTG 118  
QY 2435 GGCTTAAATTCATGTTAGTGTAGCGGCCACTGCCCTTTGTCAAAATGTGATGCACAT 2494  
Db 117 GGCTTAAATTCATGTTAGTGTAGCGGCCACTGCCCTTTGTCAAAATGTGATGCACAT 58  
QY 2495 AAGCAGTATCCAGCATGAATGTAATTTACTTGAAGTAACCTTTGGAAAAAGAAAT 2550  
Db 57 AAGCAGTATCCAGCATGAATGTAATTTACTTGAAGTAACCTTTGGAAAAAGAAAT 2

RESULT 14  
BI457843  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE

BI457843 828 bp mRNA linear EST 21-AUG-2001  
603198215F1 NIH\_MGC\_96 Homo sapiens cdna clone IMAGE:527867 5',  
mRNA sequence.  
BI457843  
BI457843.1 GI:15248499  
EST.  
Homo sapiens (human)

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 828)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Pietro Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11701 row: m column: 04
High quality sequence stop: 803.

FEATURES
    source
        1..828
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5277867"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_96"
            /notes="Organ: Brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI (gtcgag
            ); Oligo-dT primed using primer 5'-TTTTTTTTTTTNN-3',
            size-selected for average insert size 2.3 kb and
            normalized to 10^6. This is a primary library enriched
            for full-length clones and constructed using the
            Cap-trapper method (Carninci, in preparation). Library
            constructed by M. Brownstein (NIH/NHGRI, National
            Institutes of Health). Note: this is a NIH_MGC Library."
            276 a 155 c 222 g 165 t

BASE COUNT
    276 a 155 c 222 g 165 t

Query Match
    Best Local Similarity 20.7%; Score 766; DB 12; Length 828;
    Matches 808; Conservative 0; Mismatches 15; Indels 3; Gaps 3;

1684 GGAAGGAGAGACAGCCAGCTGAAACCAAGAGCTGAGGATGCGTCCACCCCCAC 1743
    Db 4 GGAGTGGGAGAGCAAGCCAGCTGAAACCAAGAGCTGAGGATGCGTCCACCCCCAC 63

1744 TGAACCTGGAGCTTCAGGAAGTGGTGGAGGAGGATCAAAATGTCAAAGCCCAATG 1803
    Db 64 TGAACCTGGAGGATTCAGGAAGTGGTGGAGGAGGATCAAAATGTCAAAGCCCAATG 123

1804 GCCTCTGAGAGCAAAATCAGCAAGCCCGAAGTCTCTGAGGATGTCGATCTAGATCTGAA 1863
    Db 124 GCCTCTGAGAGCAAAATCAGCAAGCCCGAAGTCTCTGAGGATGTCGATCTAGATCTGAA 183

1864 GAAGCTAAGACGATCTCTTCTTCACTGAAGGAAAGAACGCCGCCATTCACTGTAGCAGCTTC 1923
    Db 184 GAAGCTAAGACGATCTCTTCTTCACTGAAGGAAAGAACGCCGCCATTCACTGTAGCAGCTTC 243

1924 ATTTCAAAGCACTCTGTCTGAAGGCCCAAAACTGTGTCCCACTATCAGGAAAGGCTG 1983
    Db 244 ATTTCAAAGCACTCTGTCTGAAGGCCCAAAACTGTGTCCCACTATCAGGAAAGGCTG 303

1984 GAGCATGTGAGAGCAGAGTCAGAGTCTCTGGGTGGAAGAGTGCAGAAAGAAACAAGT 2043
    Db 304 GAGTATGTGAGAGCAGAGTCAGAGTCTCTGGGTGGAAGAGTGCAGAAAGAAACAAGT 363

2044 GGAATAATGCCAAGGCTTCTTAAGAGAAATGGGAATGTGGGAAAAACAACCTGCCAAAAA 2103
    Db 364 GGAATAATGCCAAGGCTTCTTAAGAGAAATGGGAATGTGGGAAAAACAACCTGCCAAAAA 423

2104 AGAATCTAAGGAGACGAGGAGAGAGAGTAAAGGAAGTCTAGTTGGAGATGGAGAA 2163
```

```
Db 424 AGAATCTACAGAGAGACAGGAGAGAGTAAAGAAAGTTCATAGTTTGGAGATGGAGAA 483
QY 2164 TGAGAAATCTTTAGAAAATGTGTGAGACTCCGATGAA-GATGATAACAGCTCTCTCAAC 2222
Db 484 TGAGAAATCTTTAGAAAATGTGTGAGACTCCGATGAAAGATGATGATGATGATGATGAT 543
QY 2223 AACAATCTCCCAAGAACCCCAAGTCTCTGAATGGTGCAGTCTTGTAGACAAACCTTTG 2282
Db 544 AACAATCTCCCAAGAACCCCAAGTCTCTGAATGGTGCAGTCTTGTAGACAAACCTTTG 603
QY 2283 CTGAAGAATCTCACTACTCAGAAATCCAGGATGTCGAAATCCGAGTCTGGAGGGGAGAG 2342
Db 604 CTGAAGAATCTCACTACTCAGAAATCCAGGATGTCGAAATCCGAGTCTGGAGGGGAGAG 663
QY 2343 TGGTCAAAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATGAGGATG 2402
Db 664 TGGTCAAAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATGAGGATG 723
QY 2403 AGGATGAAGAGTGACAAATTCGAATGATGCTGGGCCCTTAAATTCATGTTAGTTAGCGA 2462
Db 724 AGGATGAAGAGTGACAAATTCGAATGATGCTGGGCCCTTAAATTCATGTTAGTTAGCGA 782
QY 2463 GCCACTGCCCTTTGT-CAAAATGTGATGCACATAGCAGGATATCCC 2507
Db 783 GCCACTGCCCTTTGTCCAAATGTGATGCCCATAGCGGGTATCCC 828

RESULT 15
BG676749 781 bp mRNA linear EST 01-MAY-2003
LOCUS BG676749
DEFINITION mRNA sequence.
ACCESSION BG676749
VERSION BG676749.1 GI:139008146
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 781)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10599 row: p column: 03
High quality sequence stop: 781.

FEATURES
    source
        1..781
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4748402"
            /tissue_type="squamous cell carcinoma"
            /lab_host="DH10B (T1 phage-resistant)"
            /clone_lib="NCI_CGAP_Skn4"
            /note="Organ: skin; Vector: pCMV-SpORF6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally, primer: Oligo dT.
            Average insert size 1.5kb. Library constructed by Life
            Technologies. Note: this is a NCI_CGAP Library."
            239 a 142 c 132 g 268 t

Query Match
    Best Local Similarity 20.6%; Score 764.2; DB 10; Length 781;
    99.5%; Pred. No. 9.1e-89;
```

Matches 777; Conservative 0; Mismatches 3; Indels 1; Gaps 1;			
Qy	2652	CATAAGTCTCTAAAGGCTTGTAAGCTGGGAAATATTCACCTGATAATAGCCAGATTCT	2711
Db	1	CATRAAGTCTCTAAGGCTTGTAAGCTGGGAAATATTCACCTGATAATAGCCAGATTCT	60
Qy	2712	ACTGTATTCCCAAAAGGCAATATTAAAGTAGATAGATTAGTAGTATATTGTTACACA	2771
Db	61	ACTGTATTCCCAAAAGGCAATATTAAAGTAGACAGATTAGTAGTATATTGTTACACA	120
Qy	2772	CTATTTTCGGAATTAGAGAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAAT	2831
Db	121	CTATTTTCGGAATTAGAGAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAAT	180
Qy	2832	GCACCTTTAGTATAAAGGSCACAGTTTGATATATTTTAAATGAATACCAATTTAAATTTTT	2891
Db	181	GCACCTTTAGTATAAAGGSCACAGTTTGATATATTTTAAATGAATACCAATTTAAATTTTT	240
Qy	2892	AGTATTTACCTGTTAAGAGATTATTAGTCTCTTAAATTTTTTAGGTTAAATTTCTTGCTG	2951
Db	241	AGTATTTACCTGTTAAGAGATTATTAGTCTCTTAAATTTTTTAGGTTAAATTTCTTGCTG	300
Qy	2952	TGATATATATCAGGAATTTACTATTTTATGTCCTGCTCTTAAACTACATCCTGAACTCG	3011
Db	301	TGATATATATCAGGAATTTACTATTTTATGTCCTGCTCTTAAACTACATCCTGAACTCG	360
Qy	3012	ACGTCCTCAGGTATAATAACAACAGAGCACTTTTGGGCAATTGAAAAACCAACCTACAC	3071
Db	361	ACGTCCTCAGGTATAATAACAACAGAGCACTTTTGGGCAATTGAAAAACCAACCTACAC	420
Qy	3072	TCTTCGGTGCTTAGAGAGATCTGCTGCTCCCAAATAAGCTTTTGTATCTGCCAGTGAAT	3131
Db	421	TCTTCGGTGCTTAGAGAGATCTGCTGCTCCCAAATAAGCTTTTGTATCTGCCAGTGAAT	480
Qy	3132	TTACTGTAATCAATGATTGCTTTCTTTTCTGGTGATATCTGCTTCTCATTAATTACT	3191
Db	481	TTACCGTACTCCAAATGATTGCTTTCTTTTCTGGTGATATCTGCTTCTCATTAATTACT	540
Qy	3192	GAAGCTGCAATATTTTAGTAAATACCTTCGGGATCAGTCCCCCATCTCCCGTGTAGA	3251
Db	541	GAAGCTGCAATATTTTAGTAAATACCTTCGGGATCAGTCCCCCATCTCCCGTGTAGA	600
Qy	3252	GCAAAAGTGAAGAGTTTAAAGGAGGAGAAAGAAAGAACTGTCTTACACCACTTGAGCTCAG	3311
Db	601	GCAAAAGTGAAGAGTTTAAAGGAGGAGAAAGAAAGAACTGTCTTACACCACTTGAGCTCAG	660
Qy	3312	ACCTCTAAACCTGTATTTCCCTTATGATGTCCTTTTGGAGACACTAATTTTAAATA	3371
Db	661	ACCTCTAAACCTGTATTTCCCTTATGATGTCCTTTTGGAGACACTAATTTTAAATA	720
Qy	3372	CTTACTAGCTCTGAAATATA-TTGATTTTATCAGTATTCAGGGTGAATTAACC	3430
Db	721	CTTACTAGCTCTGAAATATACTTGATTTTATCAGTATTCAGGGTGAATTAACC	780
Qy	3431	A 3431	
Db	781	A 781	

Search completed: January 6, 2004, 15:54:30  
Job time : 7194 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 09:51:46 ; Search time 199 Seconds  
(without alignments)  
8217.711 Million cell updates/sec

Title: US-09-890-549-16

Perfect score: 3705

Sequence: 1 ggccgcaggagcagtaggtg.....atttgaaaaaaaaaaaaa 3705

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*

- 1: /cgn2\_6/prodata/1/ina/5A-COMB.seq: \*
- 2: /cgn2\_6/prodata/1/ina/5B-COMB.seq: \*
- 3: /cgn2\_6/prodata/1/ina/6A-COMB.seq: \*
- 4: /cgn2\_6/prodata/1/ina/6B-COMB.seq: \*
- 5: /cgn2\_6/prodata/1/ina/PCTUS-COMB.seq: \*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2797.6	75.5	2305	4	US-09-620-312D-45
2	362	9.8	586	3	US-09-385-982-139
3	352.6	9.5	413	4	US-09-016-434-680
4	277.4	7.5	279	4	US-09-016-434-493
5	224	6.0	238	4	US-09-016-434-181
6	126.4	3.4	710	4	US-09-016-434-701
7	118.6	3.2	654	4	US-09-016-434-994
8	118.6	3.2	1638	4	US-09-484-970B-20
9	80.6	2.2	7218	1	US-08-232-463-14
10	71.6	1.9	988	4	US-09-282-146-1
11	67.8	1.8	1015	3	US-08-934-627B-1
12	54.8	1.5	10640	4	US-09-417-485D-5
13	49.2	1.3	615	3	US-08-998-416-186
14	48.4	1.3	636	3	US-08-998-416-1137
15	47.4	1.3	3057	4	US-09-601-198-55
16	47.4	1.3	6243	2	US-09-056-075-1
17	46	1.2	837	3	US-08-998-416-288
18	44	1.2	289	3	US-09-007-005-17
19	44	1.2	289	3	US-09-244-796-17
20	44	1.2	665	2	US-08-883-795A-36
21	43.8	1.2	4673	1	US-07-638-431-1
22	43.8	1.2	4673	5	PCT-US92-00018-1
23	43.8	1.2	1664976	4	US-08-916-421B-1
24	43	1.2	827	3	US-08-998-416-535
25	43	1.2	3466	1	US-08-468-036-38
26	43	1.2	3466	2	US-08-376-843-38
27	43	1.2	6124	4	US-08-213-419B-3

C 28 42.6 1.1 53332 4 US-09-801-861-3 Sequence 3, Appli  
C 29 42.2 1.1 576 4 US-09-601-198-5 Sequence 5, Appli  
30 42.2 1.1 696 3 US-09-461-697-193 Sequence 193, App  
31 42.2 1.1 699 3 US-09-461-697-191 Sequence 191, App  
32 42.2 1.1 717 3 US-09-461-697-189 Sequence 189, App  
33 42.2 1.1 774 3 US-09-461-697-187 Sequence 187, App  
34 42.2 1.1 819 3 US-09-461-697-185 Sequence 185, App  
35 42.2 1.1 1669 3 US-09-461-697-184 Sequence 184, App  
36 42.2 1.1 20674 4 US-09-641-638-651 Sequence 651, App  
37 42 1.1 7218 1 US-08-232-463-14 Sequence 14, Appl  
38 42 1.1 19124 2 US-08-487-826B-13 Sequence 13, Appl  
39 41.8 1.1 3252 3 US-09-118-442-1 Sequence 1, Appli  
40 41.8 1.1 3252 3 US-09-677-064-1 Sequence 1, Appli  
C 41 41.4 1.1 1020 4 US-09-328-475C-43 Sequence 43, Appl  
42 41.4 1.1 2169 4 US-09-434-408-3 Sequence 3, Appli  
43 41.4 1.1 26000 4 US-09-843-376-10 Sequence 10, Appl  
44 41 1.1 663 3 US-08-998-416-191 Sequence 191, App  
45 41 1.1 1357 4 US-09-461-325-55 Sequence 55, Appl

#### ALIGNMENTS

RESULT 1  
US-09-620-312D-45  
; Sequence 45, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 45\_  
; LENGTH: 2905  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (173)...(2452)  
US-09-620-312D-45

Query Match 75.5%; Score 2797.6; DB 4; Length 2905;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 24; Indels 10; Gaps 2;  
Qy 4 CGCAGGAGCAGTAGGCTTAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGACGCCGGG 63  
Db 40 CCGGGTCGACGATTTCTAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGACGCCGGG 99

QY 64 ACCTGTGACAGGGCTGTAGACGCGAGAGAAAGGGGCTTTTAGCGAGGTATTTTCAGT 123  
 Db |||||  
 QY 100 ACCTGTGACAGGGCTGTAGACGCGAGAGAAAGGGGCTTTTAGCGAGGTATTTTCAGT 159  
 Db |||||  
 QY 124 GTCTGTAGACAGATGGAATCATCTCCATTTAAATAGACGCAATGGACCTCACTATCAT 183  
 Db |||||  
 QY 160 GTCTGTAGACAGATGGAATCATCTCCATTTAAATAGACGCAATGGACCTCACTATCAT 219  
 Db |||||  
 QY 184 GAGGGTAACGCCCAAGAACTTTCTCTGTGCAACAAGAACTCATCGGCTATTTGGA 243  
 Db |||||  
 QY 220 GAGGGTAACGCCCAAGAACTTTCTCTGTGCAACAAGAACTCATCGGCTATTTGGA 279  
 Db |||||  
 QY 244 AATATTTCTCAAGTACCAAGAAAGCAGTGAAGAAACAAACATGAGAAAGAGAGATGA 303  
 Db |||||  
 QY 280 AATATTTCTCAAGTACCAAGAAAGCAGTGAAGAAACAAACATGAGAAAGAGAGATGA 339  
 Db |||||  
 QY 304 CACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCTGACTGTGTATAAGAAAGAGTG 363  
 Db |||||  
 QY 340 CACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCTGACTGTGTATAAGAAAGAGTG 399  
 Db |||||  
 QY 364 GGAGAACCCAGGGCTGGAGCAGAGTCTCACAGACTCTCTACGGAACAGCAGCACTGA 423  
 Db |||||  
 QY 400 GGAGAACCCAGGGCTGGAGCAGAGTCTCACAGACTCTCTACGGAACAGCAGCACTGA 459  
 Db |||||  
 QY 424 GATTAGGCACAGACAGACCATCTCTGCTGAAAGTGCACAGCCACGCTCTCTCGAGC 483  
 Db |||||  
 QY 460 GATTAGGCACAGACAGACCATCTCTGCTGAAAGTGCACAGCCACGCTCTCTCGAGC 519  
 Db |||||  
 QY 484 CAAAGCTGACCAAGAAACAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAGC 543  
 Db |||||  
 QY 520 CAAAGCTGACCAAGAAACAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAGC 579  
 Db |||||  
 QY 544 CCTGTTTCAGGTCGATATCCCAACATCAAGAGCGGTGAGATCTTAAAGACCACTCAAC 603  
 Db |||||  
 QY 580 CCTGTTTCAGGTCGATATCCCAACATCAAGAGCGGTGAGATCTTAAAGACCACTCAAC 639  
 Db |||||  
 QY 604 AGAAAGTAAAAAATGAAATTTGTCTAGAGAGATCCAGGCATGAAGTAGAAAAATCAGA 663  
 Db |||||  
 QY 640 AGAAAGTAAAAAATGAAATTTGTCTAGAGAGATCCAGGCATGAAGTAGAAAAATCAGA 699  
 Db |||||  
 QY 664 AATCAGTGAATAACACAGATGCTTCGGGCAAAATAGAAAAATATAATGTTCCGCTGAACAG 723  
 Db |||||  
 QY 700 AATCAGTGAATAACACAGATGCTTCGGGCAAAATAGAAAAATATAATGTTCCGCTGAACAG 759  
 Db |||||  
 QY 724 GCTTAAGATGATGTTAGAAAGGTGAACCACTCAAACTAAGATTCTCGGGCCCCAAG 783  
 Db |||||  
 QY 760 GCTTAAGATGATGTTAGAAAGGTGAACCACTCAAACTAAGATTCTCGGGCCCCAAG 819  
 Db |||||  
 QY 784 CCGAAGTGAAGTGAAGAGATCTCTGAAACAGCTATTCTTAGATGACCTGGAAT 843  
 Db |||||  
 QY 820 CCGAAGTGAAGTGAAGAGATCTCTGAAACAGCTATTCTTAGATGACCTGGAAT 879  
 Db |||||  
 QY 844 AGGCCAGGTGAGTTGTCTATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGAGAA 903  
 Db |||||  
 QY 880 AGGCCAGGTGAGTTGTCTATCTTCTATCTTGAAGTGAAGAAAAATGAGAGTAGAGAA 939  
 Db |||||  
 QY 904 TCTGGAACTTCCAGGCTCTCAGAAACCTCTATAAGAGATCGAATGGCCAAAGTACCAGC 963  
 Db |||||  
 QY 940 TCTGGAACTTCCAGGCTCTCAGAAACCTCTATAAGAGATCGAATGGCCAAAGTACCAGC 999  
 Db |||||  
 QY 964 AGCTGTGTCACCAAGAGCTCAACCAACTATAAATGAGCTGAAAGCCAGTGGTG 1023  
 Db |||||  
 QY 1000 AGCTGTGTCACCAAGAGCTCAACCAACTATAAATGAGCTGAAAGCCAGTGGTG 1059  
 Db |||||  
 QY 1024 CGAAATCAAAATTCATAAAATGAGCAAAAGGAGATGTGCCCCAGGTCTGAGGTCTG 1083  
 Db |||||  
 QY 1060 CGAAATCAAAATTCATAAAATGAGCAAAAGGAGATGTGCCCCAGGTCTGAGGTCTG 1119  
 Db |||||  
 QY 1084 CATCAACCATCAGAAAGGGGAAAGATTTCTGCAAAATGAGATAGCCTGGCAGTCCGTTTC 1143  
 Db |||||  
 QY 1120 CATCAACCATCAGAAAGGGGAAAGATTTCTGCAAAATGAGATAGCCTGGCAGTCCGTTTC 1179  
 Db |||||  
 QY 1144 CACCCCTGCCGAGATGACTCCCGTGACTCCCGAGGTTAAGAGTGAGGTTCAACAGCCTGT 1203  
 Db |||||

Db |||||  
 QY 1180 CACCCCTGCCGAGATGACTCCCGTGACTCCAGGTTAAGAGTGAGGTTCAACAGCCTGT 1239  
 Db |||||  
 QY 1204 CCATCCCAAGCCACTAAGTCCAGATTCAGAGACCTCCAGCTCTTTCTGAAAGTTCTCCTCC 1263  
 Db |||||  
 QY 1240 CCATCCCAAGCCACTAAGTCCAGATTCAGAGACCTCCAGCTCTTTCTGAAAGTTCTCCTCC 1299  
 Db |||||  
 QY 1264 CAAAGCAATGAAGAAGTTTCAGGCACTGCAAGAGAGACCTCGTGGAAATGTGAGAAGAC 1323  
 Db |||||  
 QY 1300 CAAAGCAATGAAGAAGTTTCAGGCACTGCAAGAGAGACCTCGTGGAAATGTGAGAAGAC 1359  
 Db |||||  
 QY 1324 AGCTATCAATGGAGCGTCTCTGCGCAACAGCAGAGTGTTCACATCAGCTGCTTCGG 1383  
 Db |||||  
 QY 1360 AGCTATCAATGGAGCGTCTCTGCGCAACAGCAGAGTGTTCACATCAGCTGCTTCGG 1419  
 Db |||||  
 QY 1384 TTGCTCTATTGCCAACACAAACTCAGTCTAGGAACATATGATCTTTACATGGAAGAT 1443  
 Db |||||  
 QY 1420 TTGCTCTATTGCCAACACAAACTCAGTCTAGGAACATATGATCTTTACATGGAAGAT 1479  
 Db |||||  
 QY 1444 CTATTGTAAAGCCTCACTTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTT 1503  
 Db |||||  
 QY 1480 CTATTGTAAAGCCTCACTTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTT 1539  
 Db |||||  
 QY 1504 TGGGCACAGCACACAAAGGATCTATGGCAAGCAAAAATGAAAAACGAAGATTTTGA 1563  
 Db |||||  
 QY 1540 TGGGCACAGCACACAAAGGATCTATGGCAAGCAAAAATGAAAAACGAAGATTTTGA 1599  
 Db |||||  
 QY 1564 GAGACAGCCAGCTTGCAATGCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGC 1623  
 Db |||||  
 QY 1600 GAGACAGCCAGCTTGCAATGCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGC 1659  
 Db |||||  
 QY 1624 CCCTATTGTAAGTGGGTGCTCTGCTGCAAGTATGGAAGCAAGGCTCTCTCTCAGCA 1683  
 Db |||||  
 QY 1660 CCCTATTGTAAGTGGGTGCTCTGCTGCAAGTATGGAAGCAAGGCTCTCTCTCAGCA 1719  
 Db |||||  
 QY 1684 GGAGAAAGAACAAAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGGCCACCCCCAC 1743  
 Db |||||  
 QY 1720 GGAGAAAGAACAAAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGGCCACCCCCAC 1779  
 Db |||||  
 QY 1744 TGAACCTTGAAGTTCAGGAAGTCCCTTGAGAGAGGGATCAAAATGTCAAAGCCCAATG 1803  
 Db |||||  
 QY 1780 TGAACCTTGAAGTTCAGGAAGTCCCTTGAGAGAGGGATCAAAATGTCAAAGCCCAATG 1839  
 Db |||||  
 QY 1804 GCCTCCTGAAGCAGAAATCAGCAGCCGCAAGTTCCTGAGGATGTCGATCTAGATCTGAA 1863  
 Db |||||  
 QY 1840 GCCTCCTGAAGCAGAAATCAGCAGCCGCAAGTTCCTGAGGATGTCGATCTAGATCTGAA 1899  
 Db |||||  
 QY 1864 GAAGCTAAGACGATCTTCTTCACTGAAGAAAGAGCCGCCATTCCTGTAGCAGCTTC 1923  
 Db |||||  
 QY 1900 GAAGCTAAGACGATCTTCTTCACTGAAGAAAGAGCCGCCATTCCTGTAGCAGCTTC 1959  
 Db |||||  
 QY 1924 ATTTCAAGACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGGAAAGGCTG 1983  
 Db |||||  
 QY 1960 ATTTCAAGACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGGAAAGGCTG 2019  
 Db |||||  
 QY 1984 GAGCATGTCAAGCAGAGTGAAGATCTGTGGTGGAAAGATTCGAGAAAGAAACAAAGT 2043  
 Db |||||  
 QY 2020 GAGCATGTCAAGCAGAGTGAAGATCTGTGGTGGAAAGATTCGAGAAAGAAACAAAGT 2079  
 Db |||||  
 QY 2044 CGAAATCCCAAGCTTCTTAAGAAAGATGGAATGTGGGAAACAACTGGCCAAACAA 2103  
 Db |||||  
 QY 2080 GGAAATGCCAAGCTTCTTAAGAAAGATGGAATGTGGGAAACAACTGGCCAAACAA 2139  
 Db |||||  
 QY 2104 AGAATCTAAAGGAGAGACAGGAAAGAAAGTAAAGAAAGTCAATAGTTTGGAGATGGAGAA 2163  
 Db |||||  
 QY 2140 AGAATCTAAAGGAGAGACAGGAAAGAAAGTAAAGAAAGTCAATAGTTTGGAGATGGAGAA 2199  
 Db |||||  
 QY 2164 TGAAATCTTGTAGAAATGTCAGACTCCGATGAAGATGATTAACAGCTTCTCTCAAA 2223  
 Db |||||  
 QY 2200 TGAAATCTTGTAGAAATGTCAGACTCCGATGAAGATGATAACAGCTTCTCTCAAA 2259  
 Db |||||  
 QY 2224 ACAATCTCCCAAGAACCCCAAGTCTCTGAATTTGTCAGGTTTGTAGACAAACCTTTGC 2283  
 Db |||||



Db 2260 ACAATCTCACAGAACCCCAAGTCTCTGAATTGGTCGAGTTTGTAGACAAACACCTTTGC 2319  
QY 2284 TGAAGAATTCACTACTCAGAAATCCAGAGATGTGGAACCTCGGAGGGGAGAAGT 2343  
Db 2320 TGAAGAATTCACTACTCAGAAATCCAGAGATGTGGAACCTCGGAGGGGAGAAGT 2379  
QY 2344 GGTCAAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAAATCGGTATTATGATGAGGATGA 2403  
Db 2380 GGTCAAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAAATCGGTATTATGATGAGGATGA 2439  
QY 2404 GGATGAAGAGTGAACAAATTCGAATGATGTGGCCCTTAAATTCATGTTAGTGTAGCGAG 2463  
Db 2440 GGATGAAGAGTGAACAAATTCGAATGATGTGGCCCTTAAATTCATGTTAGTGTAGCGAG 2499  
QY 2464 CCACTGCCCTTTGTGCAAAATGTGATCCACATAGCAGAGTATCCAGCATGAATGTAATT 2523  
Db 2500 CCACTGCCCTTTGTGCAAAATGTGATCCACATAGCAGAGTATCCAGCATGAATGTAATT 2559  
QY 2524 TACTTGGAAAGTAACTTTGGAAGAAATTCCTTCTTAAATCAAAAACAAAAA 2583  
Db 2560 TACTTGGAAAGTAACTTTGGAAGAAATTCCTTCTTAAATCAAAAACAAAAA 2612  
QY 2584 ACAAAAAACACATTTCTAAATCTAGAGATACTTTACTTTAAATTTCTTATTTAGCAGTG 2643  
Db 2613 -AAAAAACACATTTCTAAATCTAGAGATACTTTACTTTAAATTTCTTATTTAGCAGTG 2670  
QY 2644 ATGATATGATAGTCTGCTGTAAGCTTTGTAAGTGGGAAATTTCCACTGATATAGCC 2703  
Db 2671 ATGATATGATAGTCTGCTGTAAGCTTTGTAAGTGGGAAATTTCCACTGATATAG-C 2729  
QY 2704 CAGATTCTACTGATTTCCCAAAAGGCAATATTAAAGTAGATAGATTAGTAGTATATT 2763  
Db 2730 CTTTTTTTCTTTATTTCCATATGCAATATTAAAGTAGACAGATGATTAGTAGTATATT 2789  
QY 2764 GTTACACACTATTTGGAAATTAGAGAACATACAGAGGAATTTAGGGCTTAAACATTAC 2823  
Db 2790 GTTACACACTATTTGTAATTAGATATACAGAGGATTTTAGGGCTTAAACATTAC 2849  
QY 2824 GACTGAATGCATTTAGTATAAGGCAAGTGTGATATTTTAAATCAATACCA 2879  
Db 2850 GACTGAATGCATTTAGTATAAGGCAAGTGTGATATTTTAAATCAATACCA 2905

RESULT 2  
US-09-385-982-139/c  
; Sequence 139, Application US/09385982  
; Patent No. 6262334  
; GENERAL INFORMATION:  
; APPLICANT: ENDEGE, WILSON O., ET AL.  
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
; FILE REFERENCE: CCDA-260XX  
; CURRENT APPLICATION NUMBER: US/09/385,982  
; EARLIER FILING DATE: 1999-08-30  
; EARLIER APPLICATION NUMBER: 09/328,111  
; EARLIER FILING DATE: 1999-06-08  
; EARLIER APPLICATION NUMBER: 60/117,393  
; EARLIER FILING DATE: 1999-01-27  
; EARLIER APPLICATION NUMBER: 60/098,639  
; EARLIER FILING DATE: 1998-08-31  
; NUMBER OF SEQ ID NOS: 544  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 139  
; LENGTH: 586  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(586)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-139

Query Match 9.8%; Score 362; DB 3; Length 586;

Best Local Similarity 100.0%; Pred. No. 1-9e-84;  
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3137 GTACTCCAAATGATTGCTTTCTTTCTGGTGATATCTGTCTCTCATAAATTACTGAAAG 3196  
Db 363 GTACTCCAAATGATTGCTTTCTTTCTGGTGATATCTGTCTCTCATAAATTACTGAAAG 304  
QY 3197 CTGCAATATTTTAGTAATACCTTCGGGATCACTGTCCCCCATCTTCGTGTTAGAGCAAA 3256  
Db 303 CTGCAATATTTTAGTAATACCTTCGGGATCACTGTCCCCCATCTTCGTGTTAGAGCAAA 244  
QY 3257 GTGAAGAGTTTTAAAGGAGGAAGAAAGAACTGCTTTACACCACTTGAGCTCAGACTC 3316  
Db 243 GTGAAGAGTTTTAAAGGAGGAAGAAAGAACTGCTTTACACCACTTGAGCTCAGACTC 184  
QY 3317 TAAACCTGTATTTCCCTTTATGATGTCCCTTTTGTAGACACTAAATTTTAAATACTTAC 3376  
Db 183 TAAACCTGTATTTCCCTTTATGATGTCCCTTTTGTAGACACTAAATTTTAAATACTTAC 124  
QY 3377 TAGCTCTGAAATATATTGATTTTATCAGATTTTATCAGGTGAAATTTAAACCAACTAT 3436  
Db 123 TAGCTCTGAAATATATTGATTTTATCAGATTTTATCAGGTGAAATTTAAACCAACTAT 64  
QY 3437 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACTTGAGT 3496  
Db 63 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACTTGAGT 4  
QY 3497 AC 3498  
Db 3 AC 2

RESULT 3  
US-09-016-434-680  
; Sequence 680, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREWITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 680:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 413 base pairs  
; TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: HNT2NOT01  
CLONE: 269891  
US-09-016-434-680

Query Match 9.5%; Score 352.6; DB 4; Length 413;  
Best Local Similarity 93.9%; Pred. No. 4.5e-82;  
Matches 386; Conservative 0; Mismatches 23; Indels 2; Gaps 2;  
QY 1022 GGCAGAAATCAAAATTCATAAAATGAGG-CAAAAGGAGAAATGTGCCCCCAGGTCTCGAGGT 1080  
Db 3 GGCAGAAATCAAAATTCATAAAATGAGG-CAAAAGGAGAAATGTGCCCCCAGGTCTCGAGGT 62  
QY 1081 CTGCATCACCATCAGGAAGGGGAAAGATTTCTGCAAAATGAGATAGCTGGCAGTCCG 1140  
Db 63 CTGCATCACCATCAGGAAGGGGAAAGATTTCTGCAAAATGAGATAGCTGGCAGTCCG 122  
QY 1141 TTCCACCCCTGCCGAAGTACCTCCCTG-ACCTCCAGGTAGAGTGAGGTTCAACAGC 1199  
Db 123 TTCCACCCCTGCCGAAGTACCTCCCTG-ACCTCCAGGTAGAGTGAGGTTCAACAGC 182  
QY 1200 CTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCTCCAGTCTTTCTGAAAGTTCTC 1259  
Db 183 CTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCTCCAGTCTTTCTGAAAGTTCTC 242  
QY 1260 CTCCAAAGCAATGAAGAAGTTTCAGGCACCTGCAAGAGAGACCTGCTGGAAATGTCAGA 1319  
Db 243 CTCCAAAGCAATGAAGAAGTTTCAGGCACCTGCAAGAGAGACCTGCTGGAAATGTCAGA 302  
QY 1320 AGACAGTCTATCCAAAGGAGCGTCTCTTGCCCAACAGCAGGTGTTTCACATCAGTGTCT 1379  
Db 303 AGACAGTCTATCCAAAGGAGCGTCTCTTGCCCAACAGCAGGTGTTTCACATCAGTGTCT 362  
QY 1380 TCCGTGCTCTTCTTTCGCAACAACTCAGTCTAGGAACATATGCATCTT 1430  
Db 363 TCCGTGCTCTTCTTTCGCAACAACTCAGTCTAGGAACATATGCATCTT 413

## RESULT 4

US-09-016-434-493  
; Sequence 493, Application US/09016434  
; Patent No. 6500938

## GENERAL INFORMATION:

APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490

## CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HEREWITH

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

## TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 493:

SEQUENCE CHARACTERISTICS:

LENGTH: 279 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: PROSNON01

CLONE: 2272243

US-09-016-434-493

## Query Match

7.5%; Score 277.4; DB 4; Length 279;

Best Local Similarity 99.6%; Pred. No. 1.3e-62;

Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2651 GCATAAGTCTCTAAGGCTTGAAGTGGGGAATATTCACCTGTATAATAGCCAGATTC 2710  
Db 1 GCGTAAGTCTCTAAGGCTTGAAGTGGGGAATATTCACCTGTATAATAGCCAGATTC 60  
QY 2711 TACTGTATTCCTCAAAAGCAATATTAAGGTAGATAGATGATTAGTAGTATATTGTTACAC 2770  
Db 61 TACTGTATTCCTCAAAAGCAATATTAAGGTAGATAGATGATTAGTAGTATATTGTTACAC 120  
QY 2771 ACTATTTTGGAAATTAGAGAAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAA 2830  
Db 121 ACTATTTTGGAAATTAGAGAAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAA 180  
QY 2831 TGCACCTTTAGTATAAAGGGCCACAGTTTCTATATTTTAAATGAATACCAATTTAATTTT 2890  
Db 181 TGCACCTTTAGTATAAAGGGCCACAGTTTCTATATTTTAAATGAATACCAATTTAATTTT 240  
QY 2891 TAGTATTTACCTGTTAAGAGATTTTAGTCTTTAAATTT 2929  
Db 241 TAGTATTTACCTGTTAAGAGATTTTAGTCTTTAAATTT 279

## RESULT 5

US-09-016-434-181

; Sequence 181, Application US/09016434

; Patent No. 6500938

## GENERAL INFORMATION:

APPLICANT: Janice Au-Young

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING

TITLE OF INVENTION: PATHWAY GENE EXPRESSION

NUMBER OF SEQUENCES: 1490

## CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

```
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 181:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KIDNOT09
; CLONE: 1422786
; US-09-016-434-181
```

```
Query Match 6.0%; Score 224; DB 4; Length 238;
Best Local Similarity 98.3%; Pred. No. 9.1e-49;
Matches 235; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2114 GGAGAGACAGGAGAGAGAGTAAAGGAGTCTATAGTTTGGAGATGGAGAAATGAGAAATCTT 2173
Db 1 GGAGAGACAGGAGAGAGAGTAAAGGAGTCTATAGTTTGGAGATGGAGAAATGAGAAATCTT 60

QY 2174 GTAGAAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCTCAAAACAACAATCTCCA 2233
Db 61 GTAGAAAATGGTGCAGACTCCGATGAAGATGATNACAGCTTCTCAAAACAACAATCTCCA 120

QY 2234 CAAGAACCAAGTCTCTGAATTTGGTGGAGTTTGTAGACAACACCTTTGCTGAAGAATTC 2293
Db 121 CAAGAACCAAGTCTCTGAATTTGGTGGAGTTTGTAGACAACACCTTTGCTGAAGAATTC 180

QY 2294 ACTACTCAGAAATCAGAAATCCAGGATGTGAACCTCTGGGAGGAGAGTGTCAAGA 2352
Db 181 ACTACTCAGAAATCAGAAATCCAGGATGTGAACCTCTGGGAGGAGAGTGTGTCAAGA 238
```

```
RESULT 6
; US-09-016-434-701
; Sequence 701, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
```

```
; INFORMATION FOR SEQ ID NO: 701:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 710 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HEARNOT01
; CLONE: 305403
; US-09-016-434-701

Query Match 3.4%; Score 126.4; DB 4; Length 710;
Best Local Similarity 73.1%; Pred. No. 3.8e-23;
Matches 174; Conservative 1; Mismatches 62; Indels 1; Gaps 1;

QY 1313 TGTCAAGACAGTCTATCCAAATGGAGCGTCTCTTGGCCCAACAGCAGGTGTTTCACATC 1372
Db 24 TGTCAAGACAGTCTATCCAAATGGAGCGTCTCTTGGCCCAACAGCAGGTGTTTCACATC 83

QY 1373 AGCTGCTTCGGTTCCTCTATTGCAACAACAACACTCAGTCTAGG-AACATATGCATCTTT 1431
Db 84 TCCTGCTTCGGATGCCACCATTCGACAKATAAACTAAGTTTGGTAAATATGCATCACT 143

QY 1432 ACATGGAAGATCTATTGTAAGCTCCTCAATCACTCTTTAAATCTAAGGGCAACTA 1491
Db 144 TCATGCAAAATATCTGTAACCTCCTTTAAACAACTTTTCAAAATCCAAAGAAATTA 203

QY 1492 TGATGAAGGCTTTGGGCACAGACCACACAGAGGATCTATGGCAGCAAAATCAAAAC 1549
Db 204 TGATGAAGGCTTTGGACATAGCAGCATAAAGATAGATGGAAGTGCAGCAAAACCAAGC 261

RESULT 7
; US-09-016-434-994
; Sequence 994, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 994:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 base pairs
; TYPE: nucleic acid
```







QY 2691 CCTGATATACCCAGATTCTTCTGTATTCCTCCAAAAGGCAATATTAAAGGTAGATAGATGA 2750  
Db 391 AATAATAATTACAAATATTAAATAAAATCTTTTATAATAATAATAATATTATT-- 334  
QY 2751 TTAGTAGTATATTGTTACACACTATTTTGGAAATTTAGAGAACATACACAGAAGATTAGGG 2810  
Db 333 TTAATAACAATTAATAATAATAATAATAATTATTTGATAATCTTATTTAATAATTATTAA 274  
QY 2811 GCTTAAACATTACGACTGAATGCACTTTTAGTATAAAGGCGACAGTTTGTATATTTTAAA 2870  
Db 273 AGAAATAATAATATCTAATAATATTTTAACTAATTTTAAATAATTGAACATAGACTAA 214  
QY 2871 TGAATACCAATTTAATTTTGTAGTATTTTACCTGTGAAGGATTTTGTAGTCTTTTAAATTT 2930  
Db 213 ATAGTATTTCATATTAATAATATTTTATAATAATAATAATAATAATAATAATAATAATA 154  
QY 2931 TTTAGGTTAATTTCTTCTGTGATATATATAGGAATTTTACTACTTTTATGT 2982  
Db 153 GTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATT 102

## RESULT 14

US-08-998-416-1137/c

; Sequence 1137, Application US/08998416

; Patent No. 6239264

## ; GENERAL INFORMATION:

; APPLICANT: Philippsen, Peter

; APPLICANT: Pohlmann, Rainer

; APPLICANT: Steiner, Sabine

; APPLICANT: Mohr, Christine

; APPLICANT: Wendland, Jürgen

; APPLICANT: Knechtle, Philipp

; APPLICANT: Reibischung, Corinne

; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPHII

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 1152

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: No. 6239264artis Corporation

; STREET: 3054 Cornwallis Road

; CITY: Research Triangle Park

; STATE: No. 6239264th Carolina

; COUNTRY: USA

; ZIP: 27709

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; APPLICATION DATA:

; APPLICATION NUMBER: US/08/998,416

; FILING DATE: 24-DEC-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: CH 0016/97

; FILING DATE: 31-DEC-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Meigs, J. Timothy

; REGISTRATION NUMBER: 38,241

; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8587

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 1137:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 636 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; ORIGINAL SOURCE:

; ORGANISM: PAG1692RP

; US-08-998-416-1137

Query Match 1.3%; Score 48.4; DB 3; Length 636;  
Best Local Similarity 45.3%; Pred. No. 0.0068;  
Matches 214; Conservative 0; Mismatches 256; Indels 2; Gaps 1;

QY 2511 ATCAAAATGTAATTTTACTTGGAAAGTAACCTTTGGAAAGAAATTCCTTCTTAAAAATCAAAAAC 2570  
Db 571 ATTAATTAATAATTAATTTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATTC 512  
QY 2571 AAAACAAAAACACAAAAACACAACTTCAATACTAGAGATAAATTTTACTTAAATTCCTT 2630  
Db 511 AATAATATAAAAATTAATAATAAGAAATTTAAAGTTTAAAGTTTAAATTTTAAATAATAATTCCT 452  
QY 2631 CATTTTAGCAGTGATGATATGATAGCTGCTGTAGGCTTGTAACTGGGGAATATTCCA 2690  
Db 451 ATAAAAGATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 392  
QY 2691 CCGTGAATAATAGCCAGATTTCTACTGTATTTCCAAAAGGCAATATTAAAGGTAGATAGATGA 2750  
Db 391 AATAATAATTTACAATATTTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATT 334  
QY 2751 TTAGTAGTATATTGTTTACACACTATTTTGGAAATTTAGAGAACATACAGAGGAATTTAGGG 2810  
Db 333 TTAATAAAACAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 274  
QY 2811 GCTTAAACATTACGACTGAAATGCCACTTTTAGTATAAAGGCGACAGTTTGTATATTTTAAA 2870  
Db 273 AGAAATA 214  
QY 2871 TGAATACCAATTTAATTTTGTAGTATTTTACCTGTTTAAAGGATTTTGTAGTCTTTTAAATTT 2930  
Db 213 ATAGTATTTCATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 154  
QY 2931 TTTAGGTTAATTTTCTTCTGTGATATATATAGGAATTTTACTACTTTTATGT 2982  
Db 153 GTAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATT 102

## RESULT 15

US-09-601-198-55/c

; Sequence 55, Application US/09601198

; Patent No. 6531583

## ; GENERAL INFORMATION:

; APPLICANT: Caswell, Gail H.

; APPLICANT: Chen, Ellison Y.

; APPLICANT: Glass, Jennifer S.

; APPLICANT: Glass, John I.

; APPLICANT: Heiner, Cheryl R.

; APPLICANT: Leikowitz, Elliot

; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA

; FILE OF INVENTION: UREALYTICUM

; FILE REFERENCE: UAB-13452/22

; CURRENT APPLICATION NUMBER: US/09/601,198

; PRIOR FILING DATE: 2000-12-08

; PRIOR APPLICATION NUMBER: 60/073,189

; PRIOR FILING DATE: 1998-01-30

; NUMBER OF SEQ ID NOS: 181

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 55

; LENGTH: 3057

; TYPE: DNA

; ORGANISM: Ureaplasma urealyticum

; US-09-601-198-55

Query Match 1.3%; Score 47.4; DB 4; Length 3057;  
Best Local Similarity 53.5%; Pred. No. 0.028;  
Matches 99; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 2561 AATCAAAAACAAAAACAAAAACAAAAACAACTTCTAAATACTAGAGATACCTTAC 2620  
Db 534 ACCAATAAAATTCATGATTAATCCAAATAATTAATAATGTAATGCAATGGTAAACCTAA 575  
QY 2621 TTAAATTTCTTCAATTTTAGCAGTGTATGATATGCTGTAGCTGTAGGCTTGTAACTGGG 2680

Db 574 ATTACAAATTCATATTACAAATATTAAATATAAATGATCAAAATAATGAATATAAAATTAAG 515  
QY 2681 AAATATTCCACCTGATAATAGCCAGATTCTACTGTATTCCCAAAGGCAATATTAAAGT 2740  
Db 514 AATTGTAAATTTGAAATAAACCRAAATTCCTTACCTATATAAAACATAAATAATGATTG 455  
QY 2741 AGATA 2745  
Db 454 AGATA 450

Search completed: January 6, 2004, 13:54:28  
Job time : 202 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 09:57:07 ; Search time 1106 Seconds  
(without alignments)  
11594.961 Million cell updates/sec

Title: US-09-890-549-16

Perfect score: 3705

Sequence: 1 ggcgcaggagcagtagtg.....attgaaaaaaaaaaaaaa 3705

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_NA.\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3568.8	96.3	3650	11	US-09-783-732-3
2	3324.4	89.7	3543	11	US-09-783-732-1
3	2797.6	75.5	2905	13	US-10-117-722-45
4	2797.6	75.5	2905	15	US-10-037-270-45
5	2736.4	73.9	2749	9	US-09-822-849A-399
6	722	19.5	732	9	US-09-925-297-152
7	590.6	15.9	698	15	US-10-198-846-10895
8	551	14.9	565	15	US-10-060-036-4338
9	548.8	14.8	583	15	US-10-066-543-3373
10	54.4	13.9	547	15	US-10-066-543-3200
11	383.2	10.3	404	11	US-09-918-995-7225
12	374.4	10.1	410	11	US-09-918-995-7296
C 13	362	9.8	586	11	US-09-871-161-139
C 14	358	9.7	358	15	US-10-066-543-1533
C 15	279.8	7.6	283	10	US-09-998-598-1767
					Sequence 3, Appli
					Sequence 1, Appli
					Sequence 45, Appl
					Sequence 45, Appl
					Sequence 152, App
					Sequence 10895, A
					Sequence 4338, Ap
					Sequence 3200, Ap
					Sequence 7225, Ap
					Sequence 7296, Ap
					Sequence 139, App
					Sequence 1533, Ap
					Sequence 1767, Ap

#### ALIGNMENTS

##### RESULT 1

US-09-783-732-3  
; Sequence 3, Application US/09783732  
; Publication No. US20030054417A1  
; GENERAL INFORMATION:  
; APPLICANT: Chang, David D.  
; APPLICANT: Maul, Raymond S.  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: EPITHELIAL PROTEIN LOST IN NEOPLASM  
; TITLE OF INVENTION: (EPLIN)  
; FILE REFERENCE: 10809/003001  
; CURRENT APPLICATION NUMBER: US/09/783,732  
; PRIOR FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 09/658,400  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 3650  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-783-732-3

Query Match 96.3%; Score 3568.8; DB 11; Length 3650;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3627; Conservative 0; Mismatches 12; Indels 11; Gaps 4;  
QY 44 GCGCTAGGTAGAGCGCGGACCTGTGACAGGCTGGTAGCGCGCAGAGAAAGCGCG 103  
DB 9 GCGCTAGGTAGAGCGCGGACCTGTGACAGGCTGGTAGCGCGCAGAGAAAGCGCG 68  
QY 104 TTTTACCGAGGTATTTCAGTGTCTGTACACAGATGGAATCATCTCCATTTAATAGCGG 163  
DB 69 TTTTACCGAGGTATTTCAGTGTCTGTACACAGATGGAATCATCTCCATTTAATAGCGG 128  
QY 164 CAATGGACCTCACTATCATTTGAGGGTAAACAGCAAGAACTTTCTCTGTCAACAAGAAC 223  
DB 129 CAATGGACCTCACTATCATTTGAGGGTAAACAGCAAGAACTTTCTCTGTCAACAAGAAC 188

QY 224 AAGTCATCGCTATTGTGGAAATATTCTCAAGTACCAGAAAGCAGCTGAAGAAACAAAC 283  
DB 189 AAGTCATCGCTATTGTGGAAATATTCTCAAGTACCAGAAAGCAGCTGAAGAAACAAAC 248  
QY 284 ATGGAGAAGAAGAGAGTAACACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCCCTG 343  
DB 249 ATGGAGAAGAGAGTAACACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCCCTG 308  
QY 344 ACTGTGTTAAAGAAAGTGGGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGACTCT 403  
DB 309 ACTGTGTTAAAGAAAGTGGGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGACTCT 368  
QY 404 CTACGGAAACAGCAGCACTGAGATTAGGCACAGAGAGACCATCTCTCTGCTGAAGTGACA 463  
DB 369 CTACGGAAACAGCAGCACTGAGATTAGGCACAGAGAGACCATCTCTCTGCTGAAGTGACA 428  
QY 464 AGCCACGCTGCTTCTCGAGCCAAAGCTGACCAAGAGAGAACAAATCCACCCAGATCTAGA 523  
DB 429 AGCCACGCTGCTTCTCGAGCCAAAGCTGACCAAGAGAGAACAAATCCACCCAGATCTAGA 488  
QY 524 CTCAGGTCACCTCTCTGAAGCCCTCGTTACAGGCTCGATATCCCCACATCAAGGACGGTGAG 583  
DB 489 CTCAGGTCACCTCTCTGAAGCCCTCGTTACAGGCTCGATATCCCCACATCAAGGACGGTGAG 548  
QY 584 GATCTTAAAGACCACTCAACAGAAAGTAAATAATGGAAAAATTGCTAGGAAATCCAGG 643  
DB 549 GATCTTAAAGACCACTCAACAGAAAGTAAATAATGGAAAAATTGCTAGGAAATCCAGG 608  
QY 644 CATGAAGTAGAAAAATCAGAAATCACTGAAACACAGATGCTTCGGGCAAAATAGAGAAA 703  
DB 609 CATGAAGTAGAAAAATCAGAAATCACTGAAACACAGATGCTTCGGGCAAAATAGAGAAA 668  
QY 704 TATAATGTTCCGCTGAACAGGCTTAAGATGATGTTTGAGAAAGGTGAACCAACTCAAAC 763  
DB 669 TATAATGTTCCGCTGAACAGGCTTAAGATGATGTTTGAGAAAGGTGAACCAACTCAAAC 728  
QY 764 AAGATTCTCGGGCCCAAGCCGAAGTGAAGTGAAGAAAGATCTCTGAAAAACAGCTAT 823  
DB 729 AAGATTCTCGGGCCCAAGCCGAAGTGAAGTGAAGAAAGATCTCTGAAAAACAGCTAT 788  
QY 824 TCTCTAGATGACCTGGAATAGCCAGGTCAGTTGTCATCTTACATTGACTCGGAG 883  
DB 789 TCTCTAGATGACCTGGAATAGCCAGGTCAGTTGTCATCTTACATTGACTCGGAG 848  
QY 884 AAAAAATGAGAGTAGACCAAAATCTGGAACCTTCCAGCCTCTCAGAAACCTCTATAAAGAT 943  
DB 849 AAAAAATGAGAGTAGACCAAAATCTGGAACCTTCCAGCCTCTCAGAAACCTCTATAAAGAT 908  
QY 944 CGAATGGCCAAAGTACAGGCACTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAAT 1003  
DB 909 CGAATGGCCAAAGTACAGGCACTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAAT 968  
QY 1004 GAGCTGAAGCCAGTGGTGGCGAAATCAAATTCATAAATGGAGCAAAAGAGAAATGTG 1063  
DB 969 GAGCTGAAGCCAGTGGTGGCGAAATCAAATTCATAAATGGAGCAAAAGAGAAATGTG 1028  
QY 1064 CCCCCAGTCTGAGGTCTGCAATCACCCATCAGGAAGGGGAAAGATTTCTGCAATGAG 1123  
DB 1029 CCCCCAGTCTGAGGTCTGCAATCACCCATCAGGAAGGGGAAAGATTTCTGCAATGAG 1088  
QY 1124 AATAGCCTGGCAGTCGGTTCCACCCCTGCGGAAGATGACTCCC---GTGACTCCCAGGTT 1180  
DB 1089 AATAGCCTGGCAGTCGGTTCCACCCCTGCGGAAGATGACTCCCAGGTTGACTCCCAGGTT 1148  
QY 1181 AAGAGTGAGGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 1240  
DB 1149 AAGAGTGAGGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 1208  
QY 1241 AGTCTTTCTGAAGTCTCTCCCAAGCAATGAAGAGGTTTCAGGCACTTCAGAGAGAG 1300  
DB 1209 AGTCTTTCTGAAGTCTCTCTCCCAAGCAATGAAGAGGTTTCAGGCACTTCAGAGAGAG 1268

QY 1301 ACCTGCGTGGAAATGTCAAGACAGAGTCTATCCAAATGAGAGCGTCTCTTTGGCCCAACCCAGAG 1360  
DB 1269 ACCTGCGTGGAAATGTCAAGACAGAGTCTATCCAAATGAGAGCGTCTCTTTGGCCCAACCCAGAG 1328  
QY 1361 GTGTTTCAATCAGCTGCTTCCGTTGCTCTATTGCAACACCAACTCAGTCTTAGGAACA 1420  
DB 1329 GTGTTTCAATCAGCTGCTTCCGTTGCTCTATTGCAACACCAACTCAGTCTTAGGAACA 1388  
QY 1421 TATCATCTTTTACATGGAAGAATCTATTGTAAGCCTCACCTTCAATCAACTCTTTAAATCT 1480  
DB 1389 TATCATCTTTTACATGGAAGAATCTATTGTAAGCCTCACCTTCAATCAACTCTTTAAATCT 1448  
QY 1481 AAGGCCAACTATGATGAAGGCTTTGGGCACAGACCAACAAGGATCTTATGGCAAGCAAA 1540  
DB 1449 AAGGCCAACTATGATGAAGGCTTTGGGCACAGACCAACAAGGATCTTATGGCAAGCAAA 1508  
QY 1541 AATGAAAACGAGAGATTTTGGAGAGACCAAGCAGCTTGCAAAATGCAAGGGAGACCCCT 1600  
DB 1509 AATGAAAACGAGAGATTTTGGAGAGACCAAGCAGCTTGCAAAATGCAAGGGAGACCCCT 1568  
QY 1601 CACAGCCAGGGGTAGAGATGCCCTATTGCTAAAGTGGGTGCTCTGGCTGCAAGTATG 1660  
DB 1569 CACAGCCAGGGGTAGAGATGCCCTATTGCTAAAGTGGGTGCTCTGGCTGCAAGTATG 1628  
QY 1661 GAAGCCAAAGGCTCTCTCTCAGCAGGAGAGAGAACAGCCAGCTGAAAACCAAGAAAGCTG 1720  
DB 1629 GAAGCCAAAGGCTCTCTCTCAGCAGGAGAGAGAACAGCCAGCTGAAAACCAAGAAAGCTG 1688  
QY 1721 AGGATCGCTGGCCACCCCTCCTGAACTTTGGAAGTTTCAAGAGTGCCTTTGGAGAAAGG 1780  
DB 1689 AGGATCGCTGGCCACCCCTCCTGAACTTTGGAAGTTTCAAGAGTGCCTTTGGAGAAAGG 1748  
QY 1781 ATCAAAATGTCAAGCCCAAAATGGCTCTCTGAAACCAAAATCAGCAAGCCCAAGTTCCT 1840  
DB 1749 ATCAAAATGTCAAGCCCAAAATGGCTCTCTGAAACCAAAATCAGCAAGCCCAAGTTCCT 1808  
QY 1841 GAGGATGTCATCTAGATCTGAAGAACTAAGACGATCTTCTTCACTGAAGAAAGAAAGC 1900  
DB 1809 GAGGATGTCATCTAGATCTGAAGAACTAAGACGATCTTCTTCACTGAAGAAAGAAAGC 1868  
QY 1901 CGCCCATCTAGTGTGAGGCTTCAATTTCAAAGCACCTCTGTCAAGAGCCCCAAAACCTGTG 1960  
DB 1869 CGCCCATCTAGTGTGAGGCTTCAATTTCAAAGCACCTCTGTCAAGAGCCCCAAAACCTGTG 1928  
QY 1961 TCCCCACTATCAGGAAGGCTGGAGCATGTGAGCAGAGTGAAGAGTCTGTGGGTGGA 2020  
DB 1929 TCCCCACTATCAGGAAGGCTGGAGCATGTGAGCAGAGTGAAGAGTCTGTGGGTGGA 1988  
QY 2021 AGGTTGCAAGAAAGGAAACAAAGTGGAAATGCCAAGCTTCTTAAGAGAAATGGAAATGTG 2080  
DB 1989 AGGTTGCAAGAAAGGAAACAAAGTGGAAATGCCAAGCTTCTTAAGAGAAATGGAAATGTG 2048  
QY 2081 GGAAAAACAACTCTGGCAAAACAAAGAAATCTAAAGGAGAGACAGGAAAGAGAAAGTAAAGAA 2140  
DB 2049 GGAAAAACAACTCTGGCAAAACAAAGAAATCTAAAGGAGAGACAGGAAAGAGAAAGTAAAGAA 2108  
QY 2141 GTTCATAGTTTGGAGATGGAGAAATGAGAAATTTTGTAGAAATGGTGCAGACTCCGATGAA 2200  
DB 2109 GTTCATAGTTTGGAGATGGAGAAATGAGAAATTTTGTAGAAATGGTGCAGACTCCGATGAA 2168  
QY 2201 GATGATAACAGCTTCTCCAAACAACTCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCG 2260  
DB 2169 GATGATAACAGCTTCTCCAAACAACTCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCG 2228  
QY 2261 AGTTTGTAGCAACACCTTTGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCCAGGAT 2320  
DB 2229 AGTTTGTAGCAACACCTTTGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCCAGGAT 2288  
QY 2321 GTGAACTCTGGAGGAGAGTGGTCAAAGAGCTCTCTGTGGAAGAAACAGATAAAGAGA 2380  
DB 2289 GTGAACTCTGGAGGAGAGTGGTCAAAGAGCTCTCTGTGGAAGAAACAGATAAAGAGA 2348  
QY 2381 AATCGGTTATTGATGAGGATGAGGATGAAGAGTGAAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2440

Db	2349	AAATCGGTATTATGATGAGGATGAGGATGAAAGAGTGACAAAATTGCAATGATGCTGGCGCTT	2408
Qy	2441	AAATTCATGTTTAGTGTTAGCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATAAGCAG	2500
Db	2409	AAATTCATGTTAGTGTTAGCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATAAGCAG	2468
Qy	2501	GTATCCCGAGCTGAATGTAAATTACTTGGAAAGTAACCTTTGGAAGAAATTCCTTCTTAA	2560
Db	2469	GTATCCCGAGCTGAATGTAAATTACTTGGAAAGTAACCTTTGGAAGAAATTCCTTCTTAA	2528
Qy	2561	AATCAAAAACAAAACAAAACAAAACAAAACAAATCTTAATACTAGAGATAACTTTTAC	2620
Db	2529	AATCAAAAACAAAACAAAACAAAACAAAACAAATCTTAATACTAGAGATAACTTTTAC	2588
Qy	2621	TTAAAAATCTTCATTTTAGCAGTGAATGATATGCATAAGTGTGAAGCTTGTAACTGGG	2680
Db	2589	TTAAAAATCTTCATTTTAGCAGTGAATGATATGCATAAGTGTGAAGCTTGTAACTGGG	2643
Qy	2681	AAATATCCACCTGATTAATAGCCGAGTTCTACTGTATTTCCAAAAGGCAATATTAAAGT	2740
Db	2644	AAATATCCACCTGATTAATAGCCGAGTTCTACTGTATTTCCAAAAGGCAATATTAAAGT	2703
Qy	2741	AGATAGATGATTAGTAGTATATGTTACACACTATTTTGGAAATTTAGAGACATACAGAAG	2800
Db	2704	AGATAGATGATTAGTAGTATATGTTACACACTATTTTGGAAATTTAGAGACATACAGAAG	2763
Qy	2801	GAATTTAGGGCTTTAAACCAATTACGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGT	2860
Db	2764	GAATTTAGGGCTTTAAACCAATTACGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGT	2822
Qy	2861	TATTTTAAATGAATACCAATTTAAATTTTAGTATTTTACCTGTTAAGAGATTATTAGT	2920
Db	2823	TATTTTAAATGAATACCAATTTAAATTTTAGTATTTTACCTGTTAAGAGATTATTAGT	2882
Qy	2921	CTTTAAATTTTTTAGTGTAAATTTCTTCGTGTGATATATATGAGGAATTTACTACTTTAT	2980
Db	2883	CTTTAAATTTTTTAGTGTAAATTTCTTCGTGTGATATATATGAGGAATTTACTACTTTAT	2942
Qy	2981	GTCCTGCTCTCTAAACTACATCTCTGAACCTCGACGCTCTCGAGGTATATAACAACAGAGCAC	3040
Db	2943	GTCCTGCTCTCTAAACTACATCTCTGAACCTCGACGCTCTCGAGGTATATAACAACAGAGCAC	3000
Qy	3041	TTTTTGAGGCAATTTGAAAACCAACCTACACTCTTCGGTCTTAGAGAGATCTCGTGTCT	3100
Db	3001	TTTTTGAGGCAATTTGAAAACCAACCTACACTCTTCGGTCTTAGAGAGATCTCGTGTCT	3060
Qy	3101	CCCAATTAAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAATGATTCCTTTCTTT	3160
Db	3061	CCCAATTAAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAATGATTCCTTTCTTT	3120
Qy	3161	TCTGGTGATATCTGTGCTTCTCATAATTACTGAAAGCTGCAATATTTAGTAATACCTTC	3220
Db	3121	TCTGGTGATATCTGTGCTTCTCATAATTACTGAAAGCTGCAATATTTAGTAATACCTTC	3180
Qy	3221	GGGATCACTGTCCCCCATCTTCGGTGTAGAGCAAGTGAAGAGTTTAAAGGAGGAAGAA	3280
Db	3181	GGGATCACTGTCCCCCATCTTCGGTGTAGAGCAAGTGAAGAGTTTAAAGGAGGAAGAA	3240
Qy	3281	GAAGAAGCTGTCTTACACCACTTTGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGAT	3340
Db	3241	GAAGAAGCTGTCTTACACCACTTTGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGAT	3300
Qy	3341	GTCCCTTTTTTGAGACACTAAATTTTTAAATACTTACTAGCTCTGAAATATTATTGATTTT	3400
Db	3301	GTCCCTTTTTTGAGACACTAAATTTTTAAATACTTACTAGCTCTGAAATATTATTGATTTT	3360
Qy	3401	ATCACAGTATTTTCAGGCGAAATTTAAACCAACTATAGGCCCTTTTCTTTGGGATGATTTT	3460
Db	3361	ATCACAGTATTTTCAGGCGAAATTTAAACCAACTATAGGCCCTTTTCTTTGGGATGATTTT	3420
Qy	3461	CTAGCTTTAAGGTTTTGGGGACATTTAAACCTTCAGTACATTTTGTGTGACACAGTTGATAT	3520

D <sub>b</sub>	3421	CTAGTCTTTAAGGTTTGGGACATTATTAACCTTGAGTACATTTGTTGTACACAGTTGTAT	3480
Q <sub>y</sub>	3521	TCCAAATTGTATGATGGGAGGAGGTGTCCTTAAGCTGTAGGCTTTTCTTTCTACTGC	3580
D <sub>b</sub>	3481	TCCAAATTGTATGATGGGAGGAGGAGGTGTCCTTAAGCTGTAGGCTTTTCTTTCTACTGC	3540
Q <sub>y</sub>	3581	ATTATTAGAGATTAGCTTTTAATATTTTTTAGAGATGTAAACACATTCGTCTTCTTAGTC	3640
D <sub>b</sub>	3541	ATTATTAGAGATTAGCTTTTAATATTTTTTAGAGATGTAAACACATTCGTCTTCTTAGTC	3600
Q <sub>y</sub>	3641	TTACTAGTCTGAACATATTTTATTCAATAAAGATTTTAAATTAATAATTG	3690
D <sub>b</sub>	3601	TTACTAGTCTGAACATATTTTATTCAATAAAGATTTTAAATTAATAATTG	3650

## RESULT 2

```

US-09-783-732-1
; Sequence 1, Application US/09783732
; Publication No. US20030054417A1
; GENERAL INFORMATION:
; APPLICANT: Chang, David D.
; APPLICANT: Maul, Raymond S.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: EPITHELIAL PROTEIN LOST IN NEOPLASM
; TITLE OF INVENTION: (EPLIN)
; FILE REFERENCE: 16809/003001
; CURRENT APPLICATION NUMBER: US/09/783,732
; CURRENT FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 09/658,400
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-783-732-1

```

Query Match.	89.7%;	Score 3324.4;	DB 11;	Length 3543;
Best Local Similarity	99.4%;	Pred. No. 0;		
Matches 3371;	Conservative 0;	Mismatches 11;	Indels 10;	Gaps 3;
Qy	302	AACACCCGAAATCTCTCCAGCACATTAGAAAGGGACCTGACTGTGTTAAAGAGAAG	361	
Db	159	AACACCGAAATCTCTCCAGCACATTAGAAAGGGACCTGACTGTGTTAAAGAGAAG	218	
Qy	362	TGGGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGACACGACGACT	421	
Db	219	TGGGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGACACGACGACT	278	
Qy	422	GAGATTAGCACAGACAGACCATCTCTGCTGAAAGTCACAAGCCACGCTGCTTCTGGA	481	
Db	279	GAGATTAGGCACAGACGACCATCTCTGCTGAAGTCACACAGCCACGCTGCTTCTGGA	338	
Qy	482	GCCAAAGCTGACCAAGAGAACAATCCACCCAGATCTAGACTCAGGTCAGCTCCTGAA	541	
Db	339	GCCAAAGCTGACCAAGAGAACAATCCACCCAGATCTAGACTCAGGTCAGCTCCTGAA	398	
Qy	542	GCCTCGTTTCAGGGTCGATCCCCACATCAAGGACGGTGAGGATCTTAAAGACACTCA	601	
Db	399	GCCTCGTTTCAGGGTCGATCCCCACATCAAGGACGGTGAGGATCTTAAAGACACTCA	458	
Qy	602	ACAGAAAGTAAAAAATGAAAAATGTCTAGAGAAATCCAGGCATCAAGTAGAAAAATCA	661	
Db	459	ACAGAAAGTAAAAAATGAAAAATGTCTAGAGAAATCCAGGCATCAAGTAGAAAAATCA	518	
Qy	662	GAAATCAGTGAACACACAGATGCTTCGGSCAAAAATAGAGAAATATAAATGTCGGGTGAAC	721	
Db	519	GAAATCAGTGAACACACAGATGCTTCGGSCAAAAATAGAGAAATATAAATGTCGGGTGAAC	578	
Qy	722	AGGCTTAGATGATGTTTGAGAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAA	781	
Db	579	AGGCTTAGATGATGTTTGAGAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAA	638	

QY 782 AGCCGAAGTGCAGTGGAAAGGAAGATCTCTGTGAAAAAGTCTATTTCTAGATGACCTGGAA 841  
Db 639 AGCCGAAGTGCAGTGGAAAGGAAGATCTCTGTGAAAAAGTCTATTTCTAGATGACCTGGAA 698  
QY 842 ATAGGCCAGTGCAGTGTGATCTCTTCTAGATTTGACTCGGAGAAAAATAGAGTAGACGA 901  
Db 699 ATAGGCCAGTGCAGTGTGATCTCTTCTAGATTTGACTCGGAGAAAAATAGAGTAGACGA 758  
QY 902 AATCTGGAATCTCCAGCCCTCTCAGAAACCTCTATAAAGGATCGAATGGCCCAAGTACCAG 961  
Db 759 AATCTGGAATCTCCAGCCCTCTCAGAAACCTCTATAAAGGATCGAATGGCCCAAGTACCAG 818  
QY 962 GCAGCTGTGTCGCAACAAAGAGAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGT 1021  
Db 819 GCAGCTGTGTCGCAACAAAGAGAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGT 878  
QY 1022 GGGGAAATCAAAATTCATAAAATGGAGCAAAAGGAAATGTGCCCCAGGTCCTGAGGTC 1081  
Db 879 GGGGAAATCAAAATTCATAAAATGGAGCAAAAGGAAATGTGCCCCAGGTCCTGAGGTC 938  
QY 1082 TGCATCACCCTCAGGAAGGGGAAAAAGATTTCTGCAAAATGAGAATAGCCTGGCAGTCCGT 1141  
Db 939 TGCATCACCCTCAGGAAGGGGAAAAAGATTTCTGCAAAATGAGAATAGCCTGGCAGTCCGT 998  
QY 1142 TCCACCCCTGCGGAAGATGACTCCC---GTGACTCCAGGTTAAGAGTGAGGTTCAACAG 1198  
Db 999 TCCACCCCTGCGGAAGATGACTCCCAGGTGACTCCAGGTTAAGAGTGAGGTTCAACAG 1058  
QY 1199 CTGTCCATCCCAAGCCACTAGTCCAGATTCAGAGCCCTCAGTCTTTCTGAAAAGTCT 1259  
Db 1059 CTGTCCATCCCAAGCCACTAGTCCAGATTCAGAGCCCTCAGTCTTTCTGAAAAGTCT 1118  
QY 1259 CCTCCCAAGCAATGAAGATTTTCAAGCACCTGCAAGAGAGACCTGGTGGAATGTGAG 1318  
Db 1119 CCTCCCAAGCAATGAAGATTTTCAAGCACCTGCAAGAGAGACCTGGTGGAATGTGAG 1178  
QY 1319 AAGACAGTCTATCCAAATGAGGCTCTTTGGCAACAGCAGGTTTCAATCAGCTGC 1378  
Db 1179 AAGACAGTCTATCCAAATGAGGCTCTTTGGCAACAGCAGGTTTCAATCAGCTGC 1238  
QY 1379 TTCCGTTGCTCTATTGCAACCAACAACTCAGTCTAGNACATATGCAATCTTTACATGA 1438  
Db 1239 TTCCGTTGCTCTATTGCAACCAACAACTCAGTCTAGNACATATGCAATCTTTACATGA 1298  
QY 1439 AGAATCTATTGTAAGCCTCAGTCTCAATCAACTCTTTTAAATCTAAGGGCAACTATGATGA 1498  
Db 1299 AGAATCTATTGTAAGCCTCAGTCTCAATCAACTCTTTTAAATCTAAGGGCAACTATGATGA 1358  
QY 1499 GGCTTTGGGCACAGACCACCAAGGATCTATGGCAAGCAAAAATGAAAAAGAGAGATT 1558  
Db 1359 GGCTTTGGGCACAGACCACCAAGGATCTATGGCAAGCAAAAATGAAAAAGAGAGATT 1418  
QY 1559 TTGGAGAGACCCAGCTTGCAAATGCAAGGAGACCCCTCAAGCCAGGGGTAGAA 1618  
Db 1419 TTGGAGAGACCCAGCTTGCAAATGCAAGGAGACCCCTCAAGCCAGGGGTAGAA 1478  
QY 1619 GATGCCCTATTGCTAAGGTGGTGTCTGGCTGCAAGTATGGAGCCCAAGGCTCTCT 1678  
Db 1479 AATGCCCTATTGCTAAGGTGGTGTCTGGCTGCAAGTATGGAGCCCAAGGCTCTCT 1538  
QY 1679 CAGCAGGAGAGAGAACAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGGCCACCC 1738  
Db 1539 CAGCAGGAGAGAGAACAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGGCCACCC 1598  
QY 1739 CCCACTGAATTTGGAAGTTCAGGAAGTGCCTTTGGAGGAGGATCAAAATGTCAAGGCC 1798  
Db 1599 CCCACTGAATTTGGAAGTTCAGGAAGTGCCTTTGGAGGAGGATCAAAATGTCAAGGCC 1658  
QY 1799 AAATGGCTCTGTAAGAGCAATCAGAGCCCGAGTTCTCTGAGGATGTCATCTAGAT 1858  
Db 1659 AAATGGCTCTGTAAGAGCAATCAGAGCCCGAGTTCTCTGAGGATGTCATCTAGAT 1718

QY 1859 CTGAAGAAGCTTAAGACGATCTTCTTCACTGAAGAAAAAGAGCCGCCATTCTACTGTAGCA 1918  
Db 1719 CTGAAGAAGCTTAAGACGATCTTCTTCACTGAAGAAAAAGAGCCGCCATTCTACTGTAGCA 1778  
QY 1919 GCTTCATTTCAAAGCACTCTGTCAAGAGCCCAAAAACTGTGTCTCCCACTCTATCAGGAAA 1978  
Db 1779 GCTTCATTTCAAAGCACTCTGTCAAGAGCCCAAAAACTGTGTCTCCCACTCTATCAGGAAA 1838  
QY 1979 GGCTGGAGCATGTCCAGAGCAGAGTGAAGATCTGTGGTGGAGAGTTGCAGAAAGAAA 2038  
Db 1839 GGCTGGAGCATGTCCAGAGCAGAGTGAAGATCTGTGGTGGAGAGTTGCAGAAAGAAA 1898  
QY 2039 CAAGTGGAAAAATGCCAAGGCTTCTTAAGAAATGGGAATCTGGGAAAAAAACAACCTGGCAA 2098  
Db 1899 CAAGTGGAAAAATGCCAAGGCTTCTTAAGAAATGGGAATCTGGGAAAAAAACAACCTGGCAA 1958  
QY 2099 AACAAAGATCTAAGGAGAGACAGGAAAGAGTGAAGAAAGTCTAGTTTGGAGATG 2158  
Db 1959 AACAAAGATTTAAAGGAGAGACAGGAAAGAGTGAAGAAAGTCTAGTTTGGAGATG 2018  
QY 2159 GAGATGAGATCTTGTAGAAAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCCTC 2218  
Db 2019 GAGATGAGATTTTGTAGAAAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCCTC 2078  
QY 2219 AAAACAACAATCTCCACAAGAAACCAAGTCTCTGAATTTGGTCGAGTTTGTAGACAACACC 2278  
Db 2079 AAAACAACAATTTCCACAAGAAACCAAGTCTTGAATTTGGTCGAGTTTGTAGACAACACC 2138  
QY 2279 TTTGCTGAAAGATTTCACTACTCAGAAATCAGAAATCCAGGATGTGAAACTCTCGGAGGGA 2338  
Db 2139 TTTGCTGAAAGATTTCACTACTCAGAAATCAGAAATCCAGGATGTGAAACTTTGGAGGGA 2198  
QY 2339 GAAGTGTCAAAGAGCTCTCTGTGGAAAGAACAGATGAAGAGAAATCGGTATTATGATGAG 2398  
Db 2199 GAAGTGTCAAAGAGCTCTCTGTGGAAAGAACAGATGAAGAGAAATCGGTATTATGATGAG 2258  
QY 2399 GATGAGATGAAGAGTGCACAAATTCGAATCATCTCGGCCCTTAAATTCATGTAGTGTGA 2458  
Db 2259 GATGAGATGAAGAGTGCACAAATTCGAATCATCTCGGCCCTTAAATTCATGTAGTGTGA 2318  
QY 2459 GCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATGAAGCAGGTATCCCAAGCATGAAATG 2518  
Db 2319 GCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATGAAGCAGGTATCCCAAGCATGAAATG 2378  
QY 2519 TAATTTACTTTGGAAGTAACTTTGGAAAAAGAAATTCCTTCTTTAAAAATCAAAAAACAAAA 2578  
Db 2379 TAATTTACTTTGGAAGTAACTTTGGAAAAAGAAATTCCTTCTTTAAAAATCAAAAAACAAAA 2438  
QY 2579 AAAACACAAAAACACATTTCTAAATACTAGAGATAAATCTTTACTTAAATCTCTCATTTTAG 2638  
Db 2439 AAAACACAAAAACACATTTCTAAATACTAGAGATAAATCTTTACTTAAATCTCTCATTTTAG 2493  
QY 2639 CAGTGATGATATGCATAAGTGTCTGAAGGCTTGTAACTGGGAAATATTCACCTGATA 2698  
Db 2494 CAGTGATGATATGCATAAGTGTCTGAAGGCTTGTAACTGGGAAATATTCACCTGATA 2553  
QY 2699 TAGCCAGATTTCTACTGTATTTCCAAAAAGGCAATATTAAGGTAGATAGATGATTAGT 2758  
Db 2554 TAGCCAGATTTCTACTGTATTTCCAAAAAGGCAATATTAAGGTAGATAGATGATTAGT 2613  
QY 2759 ATATTGTACACACTATTTTGGAAATTAGAGAACATACAGAGGAAATTTAGGGCTTTAAAC 2818  
Db 2614 ATATTGTACACACTATTTTGGAAATTAGAGAACATACAGAGGAAATTTAGGGCTTTAAAC 2673  
QY 2819 ATTACACACTGAATGCACTTTTAGTATAAAGGGCAGCTTTGTATATTTTAAATGAATACC 2878  
Db 2674 ATTACACACTGAATGCACTTTTAGTATAAAGGGCAGCTTTGTATATTTTAAATGAATACC 2733  
QY 2879 AATTTAATTTTTTAGTATTTTACCTGTGAAGAGATTTTGTAGTCTTTTAAATTTTTTAGGTT 2938  
Db 2734 AATTTAATTTTTTAGTATTTTACCTGTGAAGAGATTTTGTAGTCTTTTAAATTTTTTAGGTT 2793  
QY 2939 AATTTCTTGTGTGATATATATGAGGAATTTTACTACTTTATGTCTCTCTCTAAACTA 2998

```
Db 2794 AATTTTCTGTGATATATAGGAATTTACTACTTTATGTCCTGCTCTCTAAACTA 2853
Qy 2999 CATCTGAACTCGAGCTCCTGAGGTATATACACAGAGCACTTTTGGAGCAATGAAA 3058
Db 2854 CATCTGAACTCGAGCTCCTGAGGTATA--ACAACAGAGCACTTTTGGAGCAATGAAA 2911
Qy 3059 AACCAACCTACACTCTTCGGTGCTTAGAGAGATCTGCTCTCCCAATAAGCTTTTGTA 3118
Db 2912 AACCAACCTACACTCTTCGGTGCTTAGAGAGATCTGCTCTCCCAATAAGCTTTTGTA 2971
Qy 3119 TCTGCCAGTGAATTTACTGTACTCCAAATGATGCTTTCTTTCTGTGTATATCTGTGCT 3178
Db 2972 TCTGCCAGTGAATTTACTGTACTCCAAATGATGCTTTCTTTCTGTGTATATCTGTGCT 3031
Qy 3179 TCTCATATTTACTGAAAGCTCGAATATTTTAGTAAATACCTTCGGGATCACTGTCCCCAT 3238
Db 3032 TCTCATATTTACTGAAAGCTCGAATATTTTAGTAAATACCTTCGGGATCACTGTCCCCAT 3091
Qy 3239 CTTCGGTGTAGAGCAAGTCAAGAGTTTAAAGGAGGAAGAAAGCACTGTCTTACAC 3298
Db 3092 CTTCGGTGTAGAGCAAGTCAAGAGTTTAAAGGAGGAAGAAAGCACTGTCTTACAC 3151
Qy 3299 CACTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGATGTCCTCTTTTGGAGAC 3358
Db 3152 CACTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGATGTCCTCTTTTGGAGAC 3211
Qy 3359 TAAATTTTAAATCTACTAGCTCGAAATATATATGATTTTATCAAGTATTTCTCAGG 3418
Db 3212 TAAATTTTAAATCTACTAGCTCGAAATATATATGATTTTATCAAGTATTTCTCAGG 3271
Qy 3419 TGAATTTAAACCACTATAGCCCTTTTCTGGATGATTTTCTAGTCTTAAAGTTTGG 3478
Db 3272 TGAATTTTAAACCACTATAGCCCTTTTCTGGATGATTTTCTAGTCTTAAAGTTTGG 3331
Qy 3479 GACATTTAAACTTTGAGTACATTTGTTGTACACAGTTGATATCCAAATTTGATGGATG 3538
Db 3332 GACATTTAAACTTTGAGTACATTTGTTGTACACAGTTGATATCCAAATTTGATGGATG 3391
Qy 3539 GAGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGACTGCAATTTATAGAGATTTAGCT 3598
Db 3392 GAGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGACTGCAATTTATAGAGATTTAGCT 3451
Qy 3599 TTAATATTTTAAAGATGTAAACATTTCTGCTTTCTAGTCTTACCTAGTCTGAAACAT 3658
Db 3452 TTAATATTTTAAAGATGTAAACATTTCTGCTTTCTAGTCTTACCTAGTCTGAAACAT 3511
Qy 3659 TTTTATTTCAATAAGATTTTAAATTTAAATTTG 3690
Db 3512 TTTTATTTCAATAAGATTTTAAATTTAAATTTG 3543

RESULT 3
US-10-117-722-45
; Sequence 45, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
```

```
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 45
; LENGTH: 2905
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(2452)
US-10-117-722-45

Query Match 75.5%; Score 2797.6; DB 13; Length 2905;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2842; Conservative 0; Mismatches 24; Indels 10; Gaps 2;

Qy 4 CGCAGGAGCAGTAGGTGTTAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGAGCGCCGG 63
Db 40 CCCGGTCCACGATTCGTAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGAGCGCCGG 99
Qy 64 ACCTGTGACAGGGCTGGTAGCGCGCAGAGGAAAGCGGCTTTTAGCCAGGTATTTCACT 123
Db 100 ACCTGTGACAGGGCTGGTAGCGCGCAGAGGAAAGCGGCTTTTAGCCAGGTATTTCACT 159
Qy 124 GTCTGTAGACAAGATGGAATCATCTCCATTTTAAAGACGGCAATGGACCTCACTATCAT 183
Db 160 GTCTGTAGACAAGATGGAATCATCTCCATTTTAAAGACGGCAATGGACCTCACTATCAT 219
Qy 184 GAGGTTAACAGCAAGAAAGAACTTTCTCTGTCAACAGAACAAAGTATCGGCTATTTGTGA 243
Db 220 GAGGTTAACAGCAAGAAAGAACTTTCTCTGTCAACAGAACAAAGTATCGGCTATTTGTGA 279
Qy 244 AATATTTCTCCAAAGTACCAGAAAGCAGCTGAAAGAAACAAACATGGAGAAAGAGAAATA 303
Db 280 AATATTTCTCCAAAGTACCAGAAAGCAGCTGAAAGAAACAAACATGGAGAAAGAGAAATA 339
Qy 304 CACCGAAATATCTCTCCAGCACTTTAGAAAGGGGACCTTGACTGTGTTAAAGAAAGAGTG 363
Db 340 CACCGAAATATCTCTCCAGCACTTTAGAAAGGGGACCTTGACTGTGTTAAAGAAAGAGTG 399
Qy 364 GGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGCTCTCTACGGAACAGCAGCACTGA 423
Db 400 GGAGAACCCAGGGCTGGGAGCAGAGTCTCACACAGCTCTCTACGGAACAGCAGCACTGA 459
Qy 424 GATTAGGCACAGCAGCAGCACTCTCTCTGCTGAAAGTGACAGCCAGCTGCTTCTGGAGC 483
Db 460 GATTAGGCACAGCAGCAGCACTCTCTCTGCTGAAAGTGACAGCCAGCTGCTTCTGGAGC 519
Qy 484 CAAAGCTGACCAAGAAAGCAAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAAGC 543
Db 520 CAAAGCTGACCAAGAAAGCAAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAAGC 579
Qy 544 CCTCTTCAGGGTCAATATCCACATCAAGGACGGTGAGGATCTTAAAGACCACTCAAC 603
Db 580 CCTCTTCAGGGTCAATATCCACATCAAGGACGGTGAGGATCTTAAAGACCACTCAAC 639
Qy 604 AGAAGTAAAGAAATGGAAGAAATTTGCTAGGAGATCCAGGATGAAGTAGAAAAATCAGA 663
Db 640 AGAAGTAAAGAAATGGAAGAAATTTGCTAGGAGATCCAGGATGAAGTAGAAAAATCAGA 699
Qy 664 AATCAGTGAAGAAACAGAGTGTTCGGGCAAAATAGAGAAATATATGTTCCGCTGAAACAG 723
Db 700 AATCAGTGAAGAAACAGAGTGTTCGGGCAAAATAGAGAAATATATGTTCCGCTGAAACAG 759
Qy 724 GCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAGATTTCTCCGGGCCCAAG 783
Db 760 GCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAGATTTCTCCGGGCCCAAG 819
Qy 784 CCGAAGTGAAGTGAAGAAAGATCTCTGAAAGACAGCTATTTCTTAGATGACCTGGAAAT 843
Db 820 CCGAAGTGAAGTGAAGAAAGATCTCTGAAAGACAGCTATTTCTTAGATGACCTGGAAAT 879
Qy 844 AGGCCAGGTGAGTTGTCATCTTCTATGTTGACTCCGAGAAAAATAGAGTAGAACAAAA 903
```

Db 880 AGCCCCAGGTCAAGTGTGTCATCTTACATTTGACTCGGAGAAAAATGAGAGTAGACGAAA 939  
Qy 904 TCTGGAACCTTCCAGCGCTCTCAGAAACCTCTATAAAGGATCGAATGCCAAGTACCAGGC 963  
Db 940 TCTGGAACCTTCCAGCGCTCTCAGAAACCTCTATAAAGGATCGAATGCCAAGTACCAGGC 999  
Qy 964 AGCTGTCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAGCCAGTCGGTG 1023  
Db 1000 AGCTGTCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAGCCAGTCGGTG 1059  
Qy 1024 CGAAATCAAAATTCATAAAATGGAGCAAAAGGAGAAATGTGCCCCCAGGCTCCTGAGGTCTG 1083  
Db 1060 CGAAATCAAAATTCATAAAATGGAGCAAAAGGAGAAATGTGCCCCCAGGCTCCTGAGGTCTG 1119  
Qy 1084 CATCACCCATCAGGAAGGGGAAAAGATTTCTGCAAAATGAGAAATAGCCTGGCAGTCGGTTC 1143  
Db 1120 CATCACCCATCAGGAAGGGGAAAAGATTTCTGCAAAATGAGAAATAGCCTGGCAGTCGGTTC 1179  
Qy 1144 CACCCCTGCCAGAGATGACTCCCGTGACTCCAGGTTAAGAGTGAGGTTCAACAGCCTGT 1203  
Db 1180 CACCCCTGCCAGAGATGACTCCCGTGACTCCAGGTTAAGAGTGAGGTTCAACAGCCTGT 1239  
Qy 1204 CCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGTCTTTCTGAAAAGTTCTCCTCC 1263  
Db 1240 CCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGTCTTTCTGAAAAGTTCTCCTCC 1299  
Qy 1264 CAAAGCAATGAAGAAGTTTCAGGCACTGCGAAGAGAGACCTGCGTGGAAATGTGAGAAGAC 1323  
Db 1300 CAAAGCAATGAAGAAGTTTCAGGCACTGCGAAGAGAGACCTGCGTGGAAATGTGAGAAGAC 1359  
Qy 1324 AGTCTATCCAAATGGAGGCTCTTTGGCCCAACAGCAGGTGTTTCACATCAGTGTCTCGG 1383  
Db 1360 AGTCTATCCAAATGGAGGCTCTTTGGCCCAACAGCAGGTGTTTCACATCAGTGTCTCGG 1419  
Qy 1384 TTGCTCCTATTGCAACAAACAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGAT 1443  
Db 1420 TTGCTCCTATTGCAACAAACAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGAT 1479  
Qy 1444 CTATTGTAGCCTCAGTCTCAATCAACTCTTTAAATCTAAGGCCAACTATGATGAAGGCTT 1503  
Db 1480 CTATTGTAGCCTCAGTCTCAATCAACTCTTTAAATCTAAGGCCAACTATGATGAAGGCTT 1539  
Qy 1504 TGGGCACAGACACACAAGATCTATGGCAGACCAAAAATGAAACGGAAGATTTTGA 1563  
Db 1540 TGGGCACAGACACACAAGATCTATGGCAGACCAAAAATGAAACGGAAGATTTTGA 1599  
Qy 1564 GAGACCAGCCAGCTTGCAAAATGCAAGGAGACCCCTCACGCCCGAGGGGTAGAGATGC 1623  
Db 1600 GAGACCAGCCAGCTTGCAAAATGCAAGGAGACCCCTCACGCCCGAGGGGTAGAGATGC 1659  
Qy 1624 CCTATTGTCTAAGTGGGTGTCTGGCTGCAAGTATGGAAGCAAGGCTCTCTCTCAGCA 1683  
Db 1660 CCTATTGTCTAAGTGGGTGTCTGGCTGCAAGTATGGAAGCAAGGCTCTCTCTCAGCA 1719  
Qy 1684 GGAGAGGAGACACAGCAGCTGAAACCAAGAGCTGAGGATGCGCTGGCCACCCCCAC 1743  
Db 1720 GGAGAGGAGACACAGCAGCTGAAACCAAGAGCTGAGGATGCGCTGGCCACCCCCAC 1779  
Qy 1744 TGAATCTGGAAGTTCAGGAAGTGCCTTGAGAGAGGATCAAAATGTCAAAGCCCAATG 1803  
Db 1780 TGAATCTGGAAGTTCAGGAAGTGCCTTGAGAGAGGATCAAAATGTCAAAGCCCAATG 1839  
Qy 1804 GCCTCTGAAGACGAAATCAGCAAGCCCGAAGTTCCTGAGGATGTCGATCTAGATCTGAA 1863  
Db 1840 GCCTCTGAAGACGAAATCAGCAAGCCCGAAGTTCCTGAGGATGTCGATCTAGATCTGAA 1899  
Qy 1864 GAAGTATAGCATCTTCTTCTGCAAGGAAAGAGCCGCCATCTACTGTAGCAGCTTC 1923  
Db 1900 GAAGTATAGCATCTTCTTCTGCAAGGAAAGAGCCGCCATCTACTGTAGCAGCTTC 1959  
Qy 1924 ATTTCAAAGCACCTGTCTCAGAGCCCAAAACCTGTGTCCCCACCTATCAGGAAGGCTG 1983  
Db 1960 ATTTCAAAGCACCTGTCTCAGAGCCCAAAACCTGTGTCCCCACCTATCAGGAAGGCTG 2019

## RESULT 4

US-10-037-270-45

; Sequence 45, Application US/10037270

; Publication No. US20030104529A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyun

; APPLICANT: Chen, Rui-hong

Qy 1984 GAGCATGTGAGAGCAGAGTGAAGAGTCTGTGGTGGAGAGTTGAGAAAAGGAAACAAAGT 2043  
Db 2020 GAGCATGTGAGAGCAGAGTGAAGAGTCTGTGGTGGAGAGTTGAGAAAAGGAAACAAAGT 2079  
Qy 2044 GGAAAAATGCCAAGGCTTCTTAAGAAATGGAATGTGGGAAAAACAACTGGCAAAACAA 2103  
Db 2080 GGAAAAATGCCAAGGCTTCTTAAGAAATGGAATGTGGGAAAAACAACTGGCAAAACAA 2139  
Qy 2104 AGAATCTTAAAGAGAGACAGGGAAGAGTAAGAAAGGTCATAGTTTGGAGATGGAGAA 2163  
Db 2140 AGAATCTTAAAGAGAGACAGGGAAGAGTAAGAAAGGTCATAGTTTGGAGATGGAGAA 2199  
Qy 2164 TGAGAAATCTTTAGAAAAATGGTCAGACTCCGATGAAGATCAACAGCTTCCTCAAAACA 2223  
Db 2200 TGAGAAATCTTTAGAAAAATGGTCAGACTCCGATGAAGATCAACAGCTTCCTCAAAACA 2259  
Qy 2224 ACAATCTCCAAGAACCCAAAGTCTCTGAATTTGGTCAGATTTTGTAGACAAACACCTTTGC 2283  
Db 2260 ACAATCTCCAAGAACCCAAAGTCTCTGAATTTGGTCAGATTTTGTAGACAAACACCTTTGC 2319  
Qy 2284 TGAAGAAATTCACACTACTCAGAAATCCAGAAATCCAGAAATCGGTATTATGATGAGGATGA 2343  
Db 2320 TGAAGAAATTCACACTACTCAGAAATCCAGAAATCCAGAAATCGGTATTATGATGAGGATGA 2379  
Qy 2344 CGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATGAGGATGA 2403  
Db 2380 CGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATGAGGATGA 2439  
Qy 2404 GGATGAAGAGTGACAAATTTGCAATGATGCTGGGCCCTTAAATTTCAATTTAGTTAGCGAG 2463  
Db 2440 GGATGAAGAGTGACAAATTTGCAATGATGCTGGGCCCTTAAATTTCAATTTAGTTAGCGAG 2499  
Qy 2464 CCATGCGCCCTTTGTCAAAAATGTGATGCACATAAGCAGGTATCCAGCATGAAATGTAATT 2523  
Db 2500 CCATGCGCCCTTTGTCAAAAATGTGATGCACATAAGCAGGTATCCAGCATGAAATGTAATT 2559  
Qy 2524 TACTTGAAGTAACTTTGGAAAAGAAATTCCTTCTTAAATTCATAATTTAGTTAGCGAG 2583  
Db 2560 TACTTGAAGTAACTTTGGAAAAGAAATTCCTTCTTAAATTCATAATTTAGTTAGCGAG 2612  
Qy 2584 AAAAAACACATTTCTAAATCTAGAGATAACTTTACTTAAATTTCTTCAATTTAGCGAGT 2643  
Db 2613 - - AAAAAACACATTTCTAAATCTAGAGATAACTTTACTTAAATTTCTTCAATTTAGCGAGT 2670  
Qy 2644 ATGATATGCATAAGTGTCTGTAAGGCTTGTAACTGGGAAAATATTCCACCTGATAATAGCC 2703  
Db 2671 ATGATATGCATAAGTGTCTGTAAGGCTTGTAACTGGGAAAATATTCCACCTGATAATAG-C 2729  
Qy 2704 CAGATTTCTACTGTATTTCCCAAAAAGGCNAATTAAGGTAGATAGATGATTAGTAGTATATT 2763  
Db 2730 CTTTCTTTCTTTTATTTCCATATTGGAATATTAAAGGTAGACAGATGATTAGTAGTATATT 2789  
Qy 2764 GTTACACACTATTTTGAATTTAGAGAACATACAGAAAGGATTTTAGGGCTTTAAACATTAC 2823  
Db 2790 GTTACACACTATTTTGAATTTAGAGATCATACAGAAAGGATTTTAGGGGCTTTAAACATTAC 2849  
Qy 2824 GACTGAATGCACCTTTAGTATAAAGGCGCACAGTTTGTATATTTTAAATGAATACCA 2879  
Db 2850 GACTGAATGCACCTTTAGTATAAAGGCGCACAGTTTGTATATTTTAAATGAATACCA 2905

APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yuning  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: Tillinghast, John  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
TITLE OF INVENTION: Poly peptides  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/10/037,270  
CURRENT FILING DATE: 2002-01-04  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1104  
SOFTWARE: pt\_FL\_genes Version 1.0  
SEQ ID NO 45\_  
LENGTH: 2905  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (173)..(2452)  
US-10-037-270-45

Query Match 75.5%; Score 2797.6; DB 15; Length 2905;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 2842; Conservative 0; Mismatches 24; Indels 10; Gaps 2;  
QY 4 CGCAGGAGCAGTGGTGTAGCAGCTTGGTCGGACAGGTGCGCTAGGTAGAGCGCGGG 63  
DB 40 CCGGGTCGAGATTTCCTAGCAGCTTGGTCGGACAGGTGCGCTAGGTAGAGCGCGGG 99  
QY 64 ACCTGTACAGGGGTGTAGCAGCGCAGAGAAAGCGGCTTTAGCCAGGTATTCAGT 123  
DB 100 ACCTGTACAGGGGTGTAGCAGCGCAGAGAAAGCGGCTTTAGCCAGGTATTCAGT 159  
QY 124 GTCTGTAGACAAGTGAATCATCTCCATTATATAGAGCGCATGACCTCATTATTT 183  
DB 160 GTCTGTAGACAAGTGAATCATCTCCATTATATAGAGCGCATGACCTCATTATTT 219  
QY 184 GAGGTAAACAGCCAAAGAACTTTCTTGTCAACAAGAACTCATCGCTATTGTGGA 243  
DB 220 GAGGTAAACAGCCAAAGAACTTTCTTGTCAACAAGAACTCATCGCTATTGTGGA 279  
QY 244 AATATTTCCAAAGTACAGAAAGCAGCTGAAGAAACAAACATGGAGAAAGAGATAA 303  
DB 280 AATATTTCCAAAGTACAGAAAGCAGCTGAAGAAACAAACATGGAGAAAGATAA 339  
QY 304 CACGAAATCTCTCCAGCAGCTTTAGAAAGGGGACCTGACTGTGTTAAAGAAAGTG 363  
DB 340 CACGAAATCTCTCCAGCAGCTTTAGAAAGGGGACCTGACTGTGTTAAAGAAAGTG 399  
QY 364 GGAGAACCCAGGGCTGGGAGCAGGTCTCACACAGACTCTCTACGGAAACAGCAGACTGA 423  
DB 400 GGAGAACCCAGGGCTGGGAGCAGGTCTCACACAGACTCTCTACGGAAACAGCAGACTGA 459  
QY 424 GATTAGCAGAGAGACCATCTCTCTCTGTAAGTGACAAGCCAGCTGCTTCTGGAGC 483  
DB 460 GATTAGCAGAGAGACCATCTCTCTCTGTAAGTGACAAGCCAGCTGCTTCTGGAGC 519  
QY 484 CAAAGCTGACCAAGAGACCAATCCACCCAGATCTTAGACTCAGTCACTCTCTGAAGC 543  
DB 520 CAAAGCTGACCAAGAGACCAATCCACCCAGATCTTAGACTCAGTCACTCTCTGAAGC 579  
QY 544 CTTCTGAGGTGCGATATCCCCACATCAAGACCGTGAGGATCTTAAGACCATCTCAAC 603  
DB

DB 580 CTTCTGAGGTGCGATATCCCCACATCAAGGACGGTGAGGATCTTTAAAGACCACTCAAC 639  
QY 604 AGAAAGTAAAAAATGGAAAAATTTGCTAGAGAAATCCAGGCATGAAGTAGAAAAATCAGA 663  
DB 640 AGAAAGTAAAAAATGGAAAAATTTGCTAGAGAAATCCAGGCATGAAGTAGAAAAATCAGA 699  
QY 664 AATCAGTGAAAAACACAGATGCTTTCGGGCAAAATAGAGAAATATAATGTTCCCGCTGAACAG 723  
DB 700 AATCAGTGAAAAACACAGATGCTTTCGGGCAAAATAGAGAAATATAATGTTCCCGCTGAACAG 759  
QY 724 GCTTAAGATGATGTTTTGAGAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAAG 783  
DB 760 GCTTAAGATGATGTTTTGAGAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAAG 819  
QY 784 CGAAGTGAAGTGAAGGAAGATCTCTGAAACAGCTATTCTCTAGATGACCTGGAAT 843  
DB 820 CGAAGTGAAGTGAAGGAAGATCTCTGAAACAGCTATTCTCTAGATGACCTGGAAT 879  
QY 844 AGSCCCAGGTGCTGCTCATCTCTTACATTTGACTCGGAGAAAAATAGAGTAGACGAAA 903  
DB 880 AGSCCCAGGTGCTGCTCATCTCTTACATTTGACTCGGAGAAAAATAGAGTAGACGAAA 939  
QY 904 TCTGAACTTCCAGCGCTCTCAGAACCTCTATAAAGGATCGAATGGCCCAAGTACCAGC 963  
DB 940 TCTGNACTTCCAGCGCTCTCAGAACCTCTATAAAGGATCGAATGGCCCAAGTACCAGC 999  
QY 964 AGCTGTGTCCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1023  
DB 1000 AGCTGTGTCCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1059  
QY 1024 CGAAATCAAAATTCATAAATGGAGCAAAAGAGAAATGTGCCCCCAGGTCTCTGAGGTCTG 1083  
DB 1060 CGAAATCAAAATTCATAAATGGAGCAAAAGAGAAATGTGCCCCCAGGTCTCTGAGGTCTG 1119  
QY 1084 CATCACCATCAGGAAGGGGAAAAAGATTCTGCAAAATGAGAAATGACCTGGCAGTCCGTTTC 1143  
DB 1120 CATCACCATCAGGAAGGGGAAAAAGATTCTGCAAAATGAGAAATGACCTGGCAGTCCGTTTC 1179  
QY 1144 CACCCCTCCGGAAGATGACTCCCGTGAATCCAGGTTTAAAGTGAAGTTCAACAGCTGT 1203  
DB 1180 CACCCCTCCGGAAGATGACTCCCGTGAATCCAGGTTTAAAGTGAAGTTCAACAGCTGT 1239  
QY 1204 CCATCCCAAGCCATTAAGTCCAGATTCCAGAGCTTCCAGTCTTTCTGAAAGTTCTCTCTCC 1263  
DB 1240 CCATCCCAAGCCATTAAGTCCAGATTCCAGAGCTTCCAGTCTTTCTGAAAGTTCTCTCTCC 1299  
QY 1264 CAAAGCAATGAAGAAAGTTTCAAGGCACTTCAAGAGACCTGCGTGAATGTCAAGAC 1323  
DB 1300 CAAAGCAATGAAGAAAGTTTCAAGGCACTTCAAGAGACCTGCGTGAATGTCAAGAC 1359  
QY 1324 AGTCTATCCAAATGGAGCGTCTTTGGCCCAACAGCAGGTGTTTCAATCAGCTGCTTCCG 1383  
DB 1360 AGTCTATCCAAATGGAGCGTCTTTGGCCCAACAGCAGGTGTTTCAATCAGCTGCTTCCG 1419  
QY 1384 TTGCTCTCTATTGCAACAACTCAGTCTAGGAACATATGCATCTTTTACATGGAAGAT 1443  
DB 1420 TTGCTCTCTATTGCAACAACTCAGTCTAGGAACATATGCATCTTTTACATGGAAGAT 1479  
QY 1444 CTAITGTAAGCCTCACTTCAATCAACTCTTTAAATCTTAAGGGCACTATGATGAAGCTT 1503  
DB 1480 CTAITGTAAGCCTCACTTCAATCAACTCTTTAAATCTTAAGGGCACTATGATGAAGCTT 1539  
QY 1504 TGGGCACAGACCAACAGGATCTATGGGCAAGCAAAATATGAAACGAAGAGATTTTGA 1563  
DB 1540 TGGGCACAGACCAACAGGATCTATGGGCAAGCAAAATATGAAACGAAGAGATTTTGA 1599  
QY 1564 GAGACCAAGCCAGCTTCAATGCAAGGGAGACCTTCAAGCCAGGGGTAGAGATGC 1623  
DB 1600 GAGACCAAGCCAGCTTCAATGCAAGGGAGACCTTCAAGCCAGGGGTAGAGATGC 1659  
QY 1624 CCCTATTGCTAAGGTGGGTGCTCTGGCTGCAAGTATGGAAGCCAAAGCCCTCTCTCAGCA 1683  
DB 1660 CCCTATTGCTAAGGTGGGTGCTCTGGCTGCAAGTATGGAAGCCAAAGCCCTCTCTCAGCA 1719



QY 1684 GGAGAGGAGAGCAAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGCCACCCCCAC 1743  
DB 1720 GGGAGAGGAGAGCAAGCCAGCTGAAACCAAGAGCTGAGGATCGCTGCCACCCCCAC 1779  
QY 1744 TGAACCTTGGAAAGTTCAAGAAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAATG 1803  
DB 1780 TGAACCTTGGAAAGTTCAAGAAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAATG 1839  
QY 1804 GCTCTCTGAAGAGCAAAATCAGCAAGCCCGAAGTTCTTGAGGATGTCGATCTAGATCTGAA 1863  
DB 1840 GCTCTCTGAAGAGCAAAATCAGCAAGCCCGAAGTTCTTGAGGATGTCGATCTAGATCTGAA 1899  
QY 1864 GAAGCTTAAGCAGATCTTCTTCACTGAAGCAAGAGCCGCCCAATTCAGCTAGCAGCTTC 1923  
DB 1900 GAAGCTTAAGCAGATCTTCTTCACTGAAGCAAGAGCCGCCCAATTCAGCTAGCAGCTTC 1959  
QY 1924 ATTTCAAAGCAGCTCTGTCAAGAGCCCAAAATCTGTCCCACTATCAGAAAGGCTG 1983  
DB 1960 ATTTCAAAGCAGCTCTGTCAAGAGCCCAAAATCTGTCCCACTATCAGAAAGGCTG 2019  
QY 1984 GAGCATGTGAGAGAGAGTGAAGAGTCTGTGGGTGGAAGAGTTGAGAAAGGAACAAGT 2043  
DB 2020 GAGCATGTGAGAGAGAGTGAAGAGTCTGTGGGTGGAAGAGTTGAGAAAGGAACAAGT 2079  
QY 2044 GGAATGCAAGAGCTTCTTAAGAAAGATGGGAATGTGGGAAACCACTGCGCAAAACAA 2103  
DB 2080 GGAATGCAAGAGCTTCTTAAGAAAGATGGGAATGTGGGAAACCACTGCGCAAAACAA 2139  
QY 2104 AGAATCTAAAGGAGAGACAGGGAAGAGTGAAGAGTCAAGATGATACAGCTTCTCTCAACA 2163  
DB 2140 AGAATCTAAAGGAGAGACAGGGAAGAGTGAAGAGTCAAGATGATACAGCTTCTCTCAACA 2199  
QY 2164 TGAGAAATCTTTAGAAAATGTTGCGAGCTCCGATGAAGATGATACAGCTTCTCTCAACA 2223  
DB 2200 TGAGAAATCTTTAGAAAATGTTGCGAGCTCCGATGAAGATGATACAGCTTCTCTCAACA 2259  
QY 2224 ACAATCTCCACAGAACCCCAAGTCTCTGAATGTTGTCGAGTTTGTAGACACACCTTTGC 2283  
DB 2260 ACAATCTCCACAGAACCCCAAGTCTCTGAATGTTGTCGAGTTTGTAGACACACCTTTGC 2319  
QY 2284 TGAAGAAATTCAGTCTCAGATCAAGAAATCCAGGATGTGGAATCTCTGGGAGGAGAGT 2343  
DB 2320 TGAAGAAATTCAGTCTCAGATCAAGAAATCCAGGATGTGGAATCTCTGGGAGGAGAGT 2379  
QY 2344 GGTCAAGAGCTCTCTGTGGAAGAACAGATAAGAGAAATCGGTATTATGATGAGGATGA 2403  
DB 2380 GGTCAAGAGCTCTCTGTGGAAGAACAGATAAGAGAAATCGGTATTATGATGAGGATGA 2439  
QY 2404 GGATGAAGAGTGACAAATTTGCAATGATCCTGGGCTTTAAATTCATGTTAGTGTAGCGAG 2463  
DB 2440 GGATGAAGAGTGACAAATTTGCAATGATCCTGGGCTTTAAATTCATGTTAGTGTAGCGAG 2499  
QY 2464 CCACTGGCCCTTTGTCAAAATGTGATGCACATAAGCAGGTATCCAGCATGAAATGTAAAT 2523  
DB 2500 CCACTGGCCCTTTGTCAAAATGTGATGCACATAAGCAGGTATCCAGCATGAAATGTAAAT 2559  
QY 2524 TACTTGAAGTAACCTTTTGGAAAGAAATTCCTTCTTAAATCAAAACCAAAACAAACA 2583  
DB 2560 TACTTGAAGTAACCTTTTGGAAAGAAATTCCTTCTTAAATCAAAACCAAAACAAACA 2612  
QY 2584 ACAAAAAACACATCTTAATTAAGTGAATACCTTTACTTAAATCTTTTACCTTTAGCAGTG 2643  
DB 2613 --AAAAACACATCTTAATTAAGTGAATACCTTTACTTAAATCTTTTACCTTTAGCAGTG 2670  
QY 2644 ATGATATGCAATAGTGTGTAAGCTTTGTAACCTGGGAAATATTTCCACCTGATAATAGCC 2703  
DB 2671 ATGATATGCAATAGTGTGTAAGCTTTGTAACCTGGGAAATATTTCCACCTGATAATAG-C 2729  
QY 2704 CAGATTCTACTGTATTTCCCAAAAGGCAATATTAAGGTAGATAGATGATTAGTAGTATATT 2763  
DB 2730 CTTTTTTTCTTTATTTCCCATATGGCAATATTAAAGGTAGACAGATGATTAGTAGTATATT 2789

QY 2764 GTTACACACTATTTTGGAAATTTAGAGAACATACAGAGAAATTTAGGGCTTTAAACATTAC 2823  
DB 2790 GTTACACACTATTTTGGAAATTTAGAGAACATACAGAGAAATTTAGGGCTTTAAACATTAC 2849  
QY 2824 GACTGAATGCACTTTAGTATATAAGGGGCACAGTTTGTATATTTTAAATGAATACCA 2879  
DB 2850 GACTGAATGCACTTTAGTATATAAGGGGCACAGTTTGTATATTTTAAATGAATACCA 2905

RESULT 5  
US-09-822-849A-399  
; Sequence 399, Application US/09822849A  
; Patent No. US20020045170A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakar  
; APPLICANT: Graham, James R.  
; APPLICANT: Genetics Institute, Inc.  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6403  
; CURRENT APPLICATION NUMBER: US/09/822,849A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/195,582  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 598  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 399  
; LENGTH: 2749  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-849A-399

Query Match 73.9%; Score 2736.4; DB 9; Length 2749;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2748; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 941 GATCGAATGCCCAAGTACCAGGAGCTGTGTCCAAAACAAAGCAGCTCAACCAACTATACA 1000  
DB 1 GATCGAATGCCCAAGTACCAGGAGCTGTGTCCAAAACAAAGCAGC.CAACCAACTATACA 60  
QY 1001 AATGAGCTGAAAGCCAGTGGTGGGAAATCAAAATTCATAAAATGAGCAAAAGAGAAAT 1060  
DB 61 AATGAGCTGAAAGCCAGTGGTGGGAAATCAAAATTCATAAAATGAGC-AAAGGAGAAAT 119  
QY 1061 GTGCCCCCAGGTCCTGAGGTCTGCATCACCATCAGGAAGGGGAAAGATTTCTGCAAAAT 1120  
DB 120 GTGCCCCCAGGTCCTGAGGTCTGCATCACCATCAGGAAGGGGAAAGATTTCTGCAAAAT 179  
QY 1121 GAGAAATAGCCTGGCAGTCCGTTTCCACCCCTGCGGAGATGACTCCCGTACTCCCGAGTT 1180  
DB 180 GAGAAATAGCCTGGCAGTCCGTTTCCACCCCTGCGGAGATGACTCCCGTACTCCCGAGTT 239  
QY 1181 AAGAGTGAAGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCCAGAGCCTCC 1240  
DB 240 AAGAGTGAAGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCCAGAGCCTCC 299  
QY 1241 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAAGTTTCAGGCACCTGCAAGAGAG 1300  
DB 300 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAAGTTTCAGGCACCTGCAAGAGAG 359  
QY 1301 ACCTCGGTGGAATGTCAGAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACAGAG 1360  
DB 360 ACCTCGGTGGAATGTCAGAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACAGAG 419  
QY 1361 GTGTTTCACATCAGCTGCTTCCGTTGCTCTTATTTGCAACAAACAACTCAGTCTAGGAA 1420  
DB 420 GTGTTTCACATCAGCTGCTTCCGTTGCTCTTATTTGCAACAAACAACTCAGTCTAGGAA 479



QY 1421 TATGCATCTTTACATGGAGAATCTATTGTAAAGCCTCACTTCAATCAACTCTTTAAATCT 1480  
DB 480 TATGCATCTTTACATGGAGAATCTATTGTAAAGCCTCACTTCAATCAACTCTTTAAATCT 539  
QY 1481 AAGGGCAACTATGATGAAGCCTTTGGGCACAGACCACAAAGGATCTATGGGCAAGCAA 1540  
DB 540 AAGGGCAACTATGATGAAGCCTTTGGGCACAGACCACAAAGGATCTATGGGCAAGCAA 599  
QY 1541 AATGAAAACGAAGATTTTGGAGAGACCAGCCAGCTTGCAAATGCAAGGGAGACCCCT 1600  
DB 600 AATGAAAACGAAGATTTTGGAGAGACCAGCCAGCTTGCAAATGCAAGGGAGACCCCT 659  
QY 1601 CACAGCCAGGGGTAGAAGATGCCCTTATTGTAAAGTGGGTGCTCTGGCTCAAGATG 1660  
DB 660 CACAGCCAGGGGTAGAAGATGCCCTTATTGTAAAGTGGGTGCTCTGGCTCAAGATG 719  
QY 1661 GAAGCCAAAGCCTCTCTCAGCAGGAGAGAGGAGCAAGCCAGCTCAAAACCAAGAGCTG 1720  
DB 720 GAAGCCAAAGCCTCTCTCAGCAGGAGAGAGGAGCAAGCCAGCTCAAAACCAAGAGCTG 779  
QY 1721 AGGATCGCTGGCCACCCCTCACTGAACTTGGAAAGTTCAGGAAGTGCCTTGGAGGAAGG 1780  
DB 780 AGGATCGCTGGCCACCCCTCACTGAACTTGGAAAGTTCAGGAAGTGCCTTGGAGGAAGG 839  
QY 1781 ATCAAAATGTCAAAGCCCAATGGCTCTCTGAAGACGAAATCAGCAAGCCGAGTTCCT 1840  
DB 840 ATCAAAATGTCAAAGCCCAATGGCTCTCTGAAGACGAAATCAGCAAGCCGAGTTCCT 899  
QY 1841 GAGGATGTCGATCTAGATCTGAAGAGCTTAAGACCATCTTCTCACTGAAGGAAGAAGC 1900  
DB 900 GAGGATGTCGATCTAGATCTGAAGAGCTTAAGACCATCTTCTCACTGAAGGAAGAAGC 959  
QY 1901 CGCCCATTCACGTGAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAACCTGTG 1960  
DB 960 CGCCCATTCACGTGAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAACCTGTG 1019  
QY 1961 TCCCAACCTATCAGAAAGGCTGGAGCATGTGACAGAGAGTGAAGATCTGTGGGTGGA 2020  
DB 1020 TCCCAACCTATCAGAAAGGCTGGAGCATGTGACAGAGAGTGAAGATCTGTGGGTGGA 1079  
QY 2021 AGAGTTGAGAAAGAAACAAGTGGAATGCGAAGGCTTCTAAGAAATGGAATGTG 2080  
DB 1080 AGAGTTGAGAAAGAAACAAGTGGAATGCGAAGGCTTCTAAGAAATGGAATGTG 1139  
QY 2081 GGAAAAACAACCTCGCAAAACAAGAAATCTAAGAGGAGAGCAGGGAAGTAAGGAA 2140  
DB 1140 GGAAAAACAACCTCGCAAAACAAGAAATCTAAGAGGAGAGCAGGGAAGTAAGGAA 1199  
QY 2141 GGTCATAGTTTGGAGATGGAGATGAGAAATCTTGTAGAAATGCTGAGACTCCGATGAA 2200  
DB 1200 GGTCATAGTTTGGAGATGGAGATGAGAAATCTTGTAGAAATGCTGAGACTCCGATGAA 1259  
QY 2201 GATGATAACAGCTTCCTCAAAACAATCTCCCAAGAACCCAAAGTCTCTGAATGGTGG 2260  
DB 1260 GATGATAACAGCTTCCTCAAAACAATCTCCCAAGAACCCAAAGTCTCTGAATGGTGG 1319  
QY 2261 AGTTTGTGAGACAACCTTTGCTGAGAGATTCCTACTCAGAAATCAGAAATCCAGGAT 2320  
DB 1320 AGTTTGTGAGACAACCTTTGCTGAGAGATTCCTACTCAGAAATCAGAAATCCAGGAT 1379  
QY 2321 GTGAACTCTGGGAGGGAGAGTGGTCAAGAGCTCTCTGGAAGACAGATTAAGAGA 2380  
DB 1380 GTGAACTCTGGGAGGGAGAGTGGTCAAGAGCTCTCTGGAAGACAGATTAAGAGA 1439  
QY 2381 AATCGGTATTATGATGAGGATGAGGATGAAGAGTGAACAAATTCGAATGATGCTGGGCCCT 2440  
DB 1440 AATCGGTATTATGATGAGGATGAGGATGAAGAGTGAACAAATTCGAATGATGCTGGGCCCT 1499  
QY 2441 AAATTCATGTAGTTTGGAGGCCACTGCGCCTTTGTCAAAATGTGATGCACATAAGCAG 2500  
DB 1500 AAATTCATGTAGTTTGGAGGCCACTGCGCCTTTGTCAAAATGTGATGCACATAAGCAG 1559  
QY 2501 GTATCCCAAGCATGAATGTAATTTACTTGAAGTAAGTTCGAAAAAGAAATTCCTTCTTAA 2560

DB 1560 GTATCCCAAGCATGAATGTAATTTACTTGGAAAGTAACCTTTGAAAAAGAAATTCCTTCTTAA 1619  
QY 2561 AATCAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2620  
DB 1620 AATCAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 1679  
QY 2621 TTAATCTCTTCATTTTAGCAGTGATGATGCAATAAGTGTGTAAGGCTTGTAACTGGGG 2680  
DB 1680 TTAATCTCTTCATTTTAGCAGTGATGATGCAATAAGTGTGTAAGGCTTGTAACTGGGG 1739  
QY 2681 AAATATTCCACCTGATAATAGCCAGAGTTCTACTGTATTCCTCCAAAGGCAATATTAGGT 2740  
DB 1740 AAATATTCCACCTGATAATAGCCAGAGTTCTACTGTATTCCTCCAAAGGCAATATTAGGT 1799  
QY 2741 AGATAGATGATTAGTAGTATATTTTACACACTATTTTGGAAATTAGAGAACATACAGAAG 2800  
DB 1800 AGATAGATGATTAGTAGTATATTTTACACACTATTTTGGAAATTAGAGAACATACAGAAG 1859  
QY 2801 GAATTTAGGGGCTTAAACATTAAGCTGAATGCACTTTAGTATAAAGGGCACAGTTTGT 2860  
DB 1860 GAATTTAGGGGCTTAAACATTAAGCTGAATGCACTTTAGTATAAAGGGCACAGTTTGT 1919  
QY 2861 TATTTTAAATGAATACCAATTTAAATTTTGTATTTTACTGTATTAAGAGATTTACTTTAGT 2920  
DB 1920 TATTTTAAATGAATACCAATTTAAATTTTGTATTTTACTGTATTAAGAGATTTACTTTAGT 1979  
QY 2921 CTTTAAATTTTAAAGTTAAATTTTCTGTGTATATATATAGGAAATTTACTACTTTAT 2980  
DB 1980 CTTTAAATTTTAAAGTTAAATTTTCTGTGTATATATATAGGAAATTTACTACTTTAT 2039  
QY 2981 GTCTGTCTCTAAACTACATCTGAACTCGAGCTCTGAGGTATATAACAACAGAGCAC 3040  
DB 2040 GTCTGTCTCTAAACTACATCTGAACTCGAGCTCTGAGGTATATAACAACAGAGCAC 2099  
QY 3041 TTTTGGAGCAATGAAAAACCAACCTACACTCTTCGGTCTTTAGAGAGATCTGCTGTCT 3100  
DB 2100 TTTTGGAGCAATGAAAAACCAACCTACACTCTTCGGTCTTTAGAGAGATCTGCTGTCT 2159  
QY 3101 CCCAAATAAGCTTTTGTATCTGCGCAGTGAATTTACTGTACTCCAAATGATTTGCTTTT 3160  
DB 2160 CCCAAATAAGCTTTTGTATCTGCGCAGTGAATTTACTGTACTCCAAATGATTTGCTTTT 2219  
QY 3161 TCTGTGTATCTGTGCTTCTCATATTTACTGAAAGCTGCAATATTTTAGTAATACCTTC 3220  
DB 2220 TCTGTGTATCTGTGCTTCTCATATTTACTGAAAGCTGCAATATTTTAGTAATACCTTC 2279  
QY 3221 GGGATCACTGTCCCCCATCTTCGGTGTAGAGCAAAAGTGAAGAGTTTAAAGGAGGAAGAA 3280  
DB 2280 GGGATCACTGTCCCCCATCTTCGGTGTAGAGCAAAAGTGAAGAGTTTAAAGGAGGAAGAA 2339  
QY 3281 GAAAGAACTGTCTTACACCCTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGAT 3340  
DB 2340 GAAAGAACTGTCTTACACCCTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGAT 2399  
QY 3341 GTCCCCCTTTTGGAGACACTAAATTTTAAATCTTACTAGCTCTGAAATATATGATTTTT 3400  
DB 2400 GTCCCCCTTTTGGAGACACTAAATTTTAAATCTTACTAGCTCTGAAATATATGATTTTT 2459  
QY 3401 ATCAGAGTATTTCTAGGGTGAATTTAAACCAACTATAGSCCTTTTCTTGGGATGATTT 3460  
DB 2460 ATCAGAGTATTTCTAGGGTGAATTTAAACCAACTATAGSCCTTTTCTTGGGATGATTT 2519  
QY 3461 CTAGTCTTTAAGTTTGGGACATTTAAACCTTGAGTACATTTGTTGACACAGTTGATAT 3520  
DB 2520 CTAGTCTTTAAGTTTGGGACATTTAAACCTTGAGTACATTTGTTGACACAGTTGATAT 2579  
QY 3521 TCCAAATTTGATGATGGAGGGAGAGGTCTTAAAGCTGAGGCTTTTCTTGTACTGC 3580  
DB 2580 TCCAAATTTGATGATGGAGGGAGAGGTCTTAAAGCTGAGGCTTTTCTTGTACTGC 2639  
QY 3581 ATTTATAGAGATTTAGCTTTAAATTTTATAGAGATGTAAACATTTCTGCTTTCTTAGTC 3640

Db 2640 ATTATATAGAGATTAGCTTTTAATATATTTTATAGAGATGAAACATCTGCTTTCTTAGTC 2699

Qy 3641 TTACTAGTCTGAACACATTTTATTCATTAAGATTTTAAATTAATTTG 3690

Db 2700 TTACTTAGTCTGAACACATTTTATTCATTAAGATTTTAAATTAATTTG 2749

RESULT 6

US-09-925-297-152

Sequence 152, Application US/09925297

Patent No. US20020081659A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA105

CURRENT APPLICATION NUMBER: US/09/925,297

PRIOR FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05989

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 928

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 152

LENGTH: 732

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: (729)

OTHER INFORMATION: n equals a.t.g, or c

US-09-925-297-152

Query Match 19.5%; Score 722; DB 9; Length 732;

Best Local Similarity 99.3%; Pred. No. 7.9e-166;

Matches 719; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1096 GGAAGGGGAAAGATTCTTGCAAATCAGATAGCCTGGCAGTCGTTCCACCCCTGCCGA 1155

Db 1 GGNAGGGGAAAGATTCTTGCAAATGAGATAGCCTGGCAGTCGTTCCACCCCTGCCGA 60

Qy 1156 AGATGACTCCCGTACTCCAGGTTAAGAGTGAAGTTCACAGCCTGTCATCCCAAGCC 1215

Db 61 AGATGACTCCCGTACTCCAGGTTAAGAGTGAAGTTCACAGCCTGTCATCCCAAGCC 120

Qy 1216 ACTAAGTCAGATTCCAGAGCCTCAGTCTTTCTGAAAGTTCCTCCCAAGCAATGAA 1275

Db 121 ACTAAGTCAGATTCCAGAGCCTCAGTCTTTCTGAAAGTTCCTCCCAAGCAATGAA 180

Qy 1276 GAAGTTTCAGGCACCTGCGAAGAGAGACCTGCGTGAATGTCAGAGACAGTCTATCCAA 1335

Db 181 GAAGTTTCAGGCACCTGCGAAGAGAGACCTGCGTGAATGTCAGAGACAGTCTATCCAA 240

Qy 1336 GGAGGCTCTCTTGGCCAACCCAGCAGGTGTTTCACATCAGCTGCTTCCGTTGCTCCTATTG 1395

Db 241 GGAGGCTCTCTTGGCCAACCCAGCAGGTGTTTCACATCAGCTGCTTCCGTTGCTCCTATTG 300

Qy 1396 CAACAACAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGATCTATTGTAAGCC 1455

Db 301 CAACAACAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGATCTATTGTAAGCC 360

Qy 1456 TCACCTTCATCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACC 1515

Db 361 TCACCTTCATCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACC 420

Qy 1516 ACACAAGGATCTATGGGCAAGCAAAATGAAAACGAGAGATTTTGGAGACCCAGGCCA 1575

Db 421 ACACAAGGATCTATGGGCAAGCAAAATGAAAACGAGAGATTTTGGAGACCCAGGCCA 480

Qy 1576 GCTTGCAATGCAAGGGAGACCCCTTCACAGCCCGGGGTAGAAGATGCCCTATTGCTAA 1635

Db 481 GCTTGCAATGCAAGGGAGACCCCTTCACAGCCCGGGGTAGAAGATGCCCTATTGCTAA 540

Qy 1636 GGTGGGTCTCTGGCTGCAAGTATGGAAGCCAAAGCCTCTCTCAGCAGGAGAGGAAGA 1695

Db 541 GGTGGGTCTCTKCTGCAAGTATGGAAGCCAAAGCCTCTCTCAGCAGGAGAGGAAGA 600

Qy 1696 CAAGCCAGCTGAACCAAGAGAGCTGAGGATCGCCACCCCACTGAACCTTGGAG 1755

Db 601 CAAGCCAGCTGAACCAAGAGAGCTGAGGATCGCCACCCCACTGAACCTTGGAG 660

Qy 1756 TTCAGGAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAAAATGCCCTCTCTGAAGA 1815

Db 661 TTCAGGAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAAAATGSGCTCTCTGAAGA 720

Qy 1816 CGAA 1819

Db 721 CGAA 724

RESULT 7

US-10-198-846-10895

Sequence 10895, Application US/10198846

Publication No. US2003009974A1

GENERAL INFORMATION:

APPLICANT: Lillie, James

APPLICANT: Xu, Yongyao

APPLICANT: Wang, Youzhen

APPLICANT: Steinmann, Kathleen

TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

TITLE OF INVENTION: THERAPY OF BREAST CANCER

FILE REFERENCE: MRI-049

CURRENT APPLICATION NUMBER: US/10/198,846

CURRENT FILING DATE: 2002-07-18

PRIOR APPLICATION NUMBER: 60/306,220

PRIOR FILING DATE: 2001-07-18

NUMBER OF SEQ ID NOS: 14084

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10895

LENGTH: 698

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

LOCATION: 1, 2, 691, 692, 693, 694, 695, 696, 697, 698

OTHER INFORMATION: n = A,T,C or G

US-10-198-846-10895

Query Match 15.9%; Score 590.6; DB 15; Length 698;

Best Local Similarity 92.8%; Pred. No. 1e-133;

Matches 649; Conservative 0; Mismatches 4; Indels 46; Gaps 1;

Qy 1 GGCGCAGGAGCAGTAGGTTTACAGCTTGTGCGACAGGTGCGCTAGGTAGAGCGCC 60

Db 38 GGCGCAGGAGCAGTAGGTTTACAGCTTGTGCGACAGGTGCGCTAGGTAGAGCGCC 97

Qy 61 GGACCTGTGACAGGCTGCTAGCAGCAGAGGAAAGCGGCTTTTAGCCAGGTATTC 120

Db 98 GGACCTGTGACAGGCTGCTAGCAGCAGAGGAAAGCGGCTTTTAGCCAGGTATTC 157

Qy 121 AGTGTCTGTAGACAAGATGGAATCATCTCCATTTAATAGACGCAATGACCTCATATC 180

Db 158 AGTGTCTGTAGACAAGATGGAATCATCTCCATTTAATAGACGCAATGACCTCATATC 217

Qy 181 ATTGAGGTTAAAGCCAAAGAACTTTCTCTTGTCAACAAGAACAGTCACTGCTATTGT 240

Db 218 ATTGAGGTTAAAGCCAAAGAACTTTCTCTTGTCAACAAGAACAGTCACTGCTATTGT 277

Qy 241 GGAATATTCTCCAGTACCAGAACCACTGGAAGAAACAAACATGGAGAAAGAGAAAG 300

Db 278 GGAATATTCTCCAA----- 292

Qy 301 TAAACCCGAAAATCTCTCCACGACCTTTAGAAAGGGGACCCCTGACTGTGTGTTAAAGAGAA 360

Db 293 -AACACCCGAAAATCTCTCCACGACCTTTAGAAAGGGGACCCCTGACTGTGTGTTAAAGAGAA 351

QY 361 GTGGGAGACCCAGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGAAACAGCAGCAC 420  
Db 352 GTGGGAGACCCAGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGAAACAGCAGCAC 411  
QY 421 TGAGATTAGGCACAGACAGACCATCTCTCTGCTGAGTGCAGAACCCAGCTGCTTCTGG 480  
Db 412 TGAGATTAGGCACAGACAGACCATCTCTCTGCTGAGTGCAGAACCCAGCTGCTTCTGG 471  
QY 481 AGCCAAAGCTGACCAAGAGAAACAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGA 540  
Db 472 AGCCAAAGCTGACCAAGAGAAACAAATCCACCCAGATCTAGACTCAGGTCACTCTCTGA 531  
QY 541 AGCCCTGTTCAAGGTGCGATATCCCAATCAAGACCGGTGAGGATCTTAAAGACCACTC 600  
Db 532 AGCCCTGTTCAAGGTGCGATATCCCAATCAAGACCGGTGAGGATCTTAAAGACCACTC 591  
QY 601 AACAGAAAGTAAATAAATGGAATAATGTCTAGGAAATCCAGGATGAAGTAGAAAAATC 660  
Db 592 AACAGAAAGTAAATAAATGGAATAATGTCTAGGAAATCCAGGATGAAGTAGAAAAATC 651  
QY 661 AGAAATCAGTGAACACACAGATGCTTCGGGCAAAATAGA 699  
Db 652 AGAAATCAGTGAACACACAGATGCTTCGGGCAAAATAGA 690

## RESULT 8

US-10-060-036-4338  
; Sequence 4338, Application US/10060036  
; Publication No. US20030073144A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Persing, David H.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Jiang, Yugu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060.036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SEQ ID NO 4338  
; SOFTWARE: FastSeq for Windows Version 4.0  
; LENGTH: 565  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 416..418, 556  
; OTHER INFORMATION: n = A,T,C or G  
US-10-060-036-4338

Query Match 14.9%; Score 551; DB 15; Length 565;  
Best Local Similarity 99.5%; Pred. No. 4.4e-124;  
Matches 562; Conservative 0; Mismatches 2; Indels 1; Gaps 1;  
QY 1925 TTCAAAGACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGAAAGGCTGG 1984  
Db 1 TTTCAAAGACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGAAAGGCTGG 60  
QY 1985 AGCATGTCAGAGCAGAGTGAAGTCTGTGGTGGAGAGTTCGAGAAAGGAAACAAGTG 2044  
Db 61 AGCATGTCAGAGCAGAGTGAAGTCTGTGGTGGAGAGTTCGAGAAAGGAAACAAGTG 120  
QY 2045 GAAATGCCAAGGCTCTAAGAGAAATGGAATGTGGGAAACACACCTGGCAAAACAA 2104  
Db 121 GAAATGCCAAGGCTCTAAGAGAAATGGAATGTGGGAAACACACCTGGCAAAACAA 180  
QY 2105 GAATCTAAGAGAGACAGCGAAGAGAAAGTGAAGAGGTCATAGTTTGAGATGGAGAT 2164  
Db 181 GAATCTAAGAGAGACAGCGAAGAGAAAGTGAAGAGGTCATAGTTTGAGATGGAGAT 240

QY 2165 GAGAAATCTGTAGAAAAATGGTGAGACTCCGATGCAAGATGAACAGCTTCTCTCAAAACA 2224  
Db 241 GAGAAATCTGTAGAAAAATGGTGAGACTCCGATGCAAGATGAACAGCTTCTCTCAAAACA 300  
QY 2225 CAATCTCCCAAGAACCCCAAGTCTCTGAATGGTCGAGTTTGTAGACAAACACCTTTGCT 2284  
Db 301 CAATCTCCCAAGAACCCCAAGTCTCTGAATGGTCGAGTTTGTAGACAAACACCTTTGCT 360  
QY 2285 GAAGAAATCTACTACTCTCAGAAATCCAGAAATCCAGGATGTGGAACCTCTGGAGGGAG-AA 2343  
Db 361 GAAGAAATCTACTACTCTCAGAAATCCAGAAATCCAGGATGTGGAACCTCTGGAGGGAGNGT 420  
QY 2344 GGTCAAAAGACTCTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTATTTATGATGAGGATGA 2403  
Db 421 GGTCAAAAGACTCTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTATTTATGATGAGGATGA 480  
QY 2404 GGATGAAGAGTGAACAAATTCGAATGATGCTGGGCCCTTAAATTCATGTTAGTGTAGCGAG 2463  
Db 481 GGATGAAGAGTGAACAAATTCGAATGATGCTGGGCCCTTAAATTCATGTTAGTGTAGCGAG 540  
QY 2464 CCACTGCCCTTTGTCAAAATGTGAT 2488  
Db 541 CCACTGCCCTTTGTCAAAATGTGAT 565

## RESULT 9

US-10-066-543-3373  
; Sequence 3373, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066.543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3373  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 534  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-3373

Query Match 14.8%; Score 548.8; DB 15; Length 583;  
Best Local Similarity 97.8%; Pred. No. 1.5e-123;  
Matches 580; Conservative 0; Mismatches 3; Indels 10; Gaps 2;  
QY 2543 AAAAGAAATCTCTTTAAATCAAAAACAAAACAAAACAAAACAAATTTCTTAAA 2602  
Db 1 AAAAGAAATCTCTTTAAATCAAAAACAAAACAAAACAAAACAAATTTCTTAAA 51  
QY 2603 TACTAGAGATAACTTTTACTTAAATCTTCTTTTAGAGAGTATGATATGCAATAGTGTG 2662  
Db 52 TACTAGAGATAACTTTTACTTAAATCTTCTTTTAGAGAGTATGATATGCAATAGTGTG 111  
QY 2663 TAAGCTTGTAACTGGGAAATATTCACCTGTATAATAGCCAGATTTCTACTGTATTTCCC 2722

Qy	3168	ATATCTGTGCTTCTCATATTAATTAAGAGCTGCAATATTTTAGTAATAACCTTCGGGATCA	3227
Db	1	ATATCTGCGCTTCTCATATTAATTAAGAGCTGCTTATTTTAGTAATAACCTTCGGGACCA	60
Qy	3228	CTGTCCCCCACTTCCGTGTTAGAGCAAGTGAAGAGTTTAAAGSGAGGAAGAAAGAA	3287
Db	61	CTGTCCCCCACTTCCGTGTTAGAGCAAGTGAAGAGTTTAAAGSGAGGAAGAAAGAA	120
Qy	3288	CTGTCTTATACACCACTTGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGATGTCCT	3347
Db	121	CTGTCTTATACACCACTTGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGATGTCCT	180

```
QY 3348 TTTTGAGACACTAATTTTAAATACCTTACTAGCTCTGAAATATATTGATTTTATACAG 3407
Db 181 TTTTGAGACACTAATTTTAAATACCTTACTAGCTCTGAAATATATTGATTTTATACAG 240
QY 3408 TATTTCTCAGGTGAAATTAACCAACTATAGGCTTTTCTTGGGATGATTTTCTAGTCT 3467
Db 241 GCTTCTCAGGTGAAATTAACCAACTATAGGCTTTTCTTGGGATGATTTTCTAGTCT 300
QY 3468 TAAAGTTTGGGGACATTTATAAATTTAGTACATATTTGTGTACACAGTTGATATCCAAAT 3527
Db 301 TAAAGTTTGGGGACATTTATAAATTTAGTACATATTTGTGTACACAGTTGATATCCACAT 360
QY 3528 TGTATGATGGGAGGAGAGGTGCTTTAAGTGTAGGCTTTTCT 3571
Db 361 TGCATGGATGGGAGGAGAGGTGCTTTAAGCTGTAGGCTCTCT 404

RESULT 12
US-09-918-995-7296
; Sequence 7296, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7296
; LENGTH: 410
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7296
```

```
Query Match 10.1%; Score 374.4; DB 11; Length 410;
Best Local Similarity 94.9%; Pred. No. 5.4e-81;
Matches 387; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 3167 GATATCTGTGCTTCTCATAATTACTGAAAGCTGCAATATTTTAGTAAATACCTTCGGGATC 3226
Db 2 GATATCTGTGCTTCTCATAATTACTGAAAGCTGCAATATTTTAGTAAATACCTTCGGGATC 61
QY 3227 ACTGTCCCCCATCTTCGTTTAGACAAAGTGAAGAGTTTAAAGGAGGAAGAAAGA 3286
Db 62 ACTGTCCCGCATGTCCGTTTAGACAAAGGAGAGCTTAAAGGAGGAAGAGATGA 121
QY 3287 ACTGTCTTACACCACTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGATGTCCTC 3346
Db 122 ACTGTCTTACACCACTTGAGCTCAGACCTCTAAACCTGTATTTCCCTTATGATGTCCTC 181
QY 3347 TTTTTCAGACACTAATTTTAAATACCTTACTAGCTCTGAAATATATTGATTTTATCACA 3406
Db 182 TTTTTCAGACACTAATTTTAAATACCTTACTAGCTCTGAAATATATTGATTTTATCACA 241
QY 3407 GTATTTCTCAGGTGAAATTAACCAACTATAGGCTTTTCTTGGGATGATTTTCTAGTC 3466
Db 242 GTATTTCTCAGGTGAAATTAACCAACTATAGGCTTTTCTTGGGATGATTTTCTAGTC 301
QY 3467 TTAAGGTTTGGGGACATTTATAAATTTAGTACATATTTGTGTACACAGTTGATATCCAAA 3526
Db 302 TTAAGGATTTGGGGACATTTATAAATTTAGTACATATGCTGGACACAGTTGATATCCAAA 361
QY 3527 TTGTATGGATGGGAGGAGGTGCTTTAAGCTGTAGGCTTTTCTTTG 3574
Db 362 TTGTATGGATGGGAGGAGATGTGCTTTAAGCTGTAGGCTATTTCTTTG 409
```

```
RESULT 13
US-09-871-161-139/c
```

```
; Sequence 139, Application US/09871161
; Publication No. US20030097666A1
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCDNA-260XX
; CURRENT APPLICATION NUMBER: US/09/871,161
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/328,111
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/117,393
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 60/098,639
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 139
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(586)
; OTHER INFORMATION: n = A,T,C or G
US-09-871-161-139

Query Match 9.8%; Score 362; DB 11; Length 586;
Best Local Similarity 100.0%; Pred. No. 7.4e-78;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3137 GTACTCCAAATGATTGCTTTCTTTCTGGTGATATCTGTGCTTCTCATAAATTAAGGAAAG 3196
Db 363 GTACTCCAAATGATTGCTTTCTTTCTGGTGATATCTGTGCTTCTCATAAATTAAGGAAAG 304
QY 3197 CTGCAATATTTTAGTAATACCTTCGGGATCACTGCCCCCATCTTCCTGTTTAGAGCAA 3256
Db 303 CTGCAATATTTTAGTAATACCTTCGGGATCACTGCCCCCATCTTCCTGTTTAGAGCAA 244
QY 3257 GTGAAGAGTTTAAAGGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3316
Db 243 GTGAAGAGTTTAAAGGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 184
QY 3317 TAAACCCCTGTATTTCCCTTATGATCTCCCTTTTATGATCTTTTATGACACTATTTTAAATACCTTAC 3376
Db 183 TAAACCCCTGTATTTCCCTTATGATCTCCCTTTTATGATCTTTTATGACACTATTTTAAATACCTTAC 124
QY 3377 TAGCTCTGAAATATATTGATTTTATCAGATTTTCTCAGGTGAAATTAACCAACTAT 3436
Db 123 TAGCTCTGAAATATATTGATTTTATCAGATTTTCTCAGGTGAAATTAACCAACTAT 64
QY 3437 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACTTTGAGT 3496
Db 63 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACTTTGAGT 4
QY 3497 AC 3498
Db 3 AC 2

RESULT 14
US-10-066-543-1533/c
; Sequence 1533, Application US/10066543
; Publication No. US20030087818A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
```

```
; APPLICANT: Smith, Carole L.
; APPLICANT: Durham, Margarita
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.563
; CURRENT APPLICATION NUMBER: US/10/066,543
; CURRENT FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 3417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1533
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-066-543-1533

Query Match
Best Local Similarity 100.0%; Score 358; DB 15; Length 358;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3139 ACTCCAAATGATGCTTTCTTTTCTGGTGATATCTGTGTTCTCTATAATTAAGT 3198
Db |||||||
QY 3199 GCAATATTTTAGTATACCTTCGGGATCACTGTCCTCCCATCTTCGGTGTAGAGCAAAAGT 3258
Db |||||||
QY 298 GCAATATTTTAGTATACCTTCGGGATCACTGTCCTCCCATCTTCGGTGTAGAGCAAAAGT 239
Db |||||||
QY 3259 GAAGAGTTTTAAAGGAGGAAGAAGAACTGCTTTACACCACTTTGAGCTCAGACCTCTA 3318
Db |||||||
QY 238 GAAGAGTTTTAAAGGAGGAAGAAGAACTGCTTTACACCACTTTGAGCTCAGACCTCTA 179
Db |||||||
QY 3319 AACCCGTGATTTCCCTTATGATGTCCTTTTGAGACACTAAATTTTAAATACTTACTA 3378
Db |||||||
QY 178 AACCCGTGATTTCCCTTATGATGTCCTTTTGAGACACTAAATTTTAAATACTTACTA 119
Db |||||||
QY 3379 GCTCTGAATATATTGATTTTATCAGACTATTTCTCAGGGTGAATTTAAACCACTATAG 3438
Db |||||||
QY 118 GCTCTGAATATATTGATTTTATCAGACTATTTCTCAGGGTGAATTTAAACCACTATAG 59
Db |||||||
QY 3439 GCCTTTTCTGGGATGATTTTCTAGTCTTAAGTTTGGGACATTAATTAAGCTTGAGT 3496
Db |||||||
58 GCCTTTTCTGGGATGATTTTCTAGTCTTAAGTTTGGGACATTAATTAAGCTTGAGT 1
```

```
RESULT 15
US-09-998-598-1767/c
; Sequence 1767, Application US/09998598
; Patent No. US2002015092A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 1767
; LENGTH: 283
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-1767

Query Match
Best Local Similarity 7.6%; Score 279.8; DB 10; Length 283;
Matches 281; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2585 CAAAAACACATTTCTAAATAGTAGTAACTTTTCTTAAATTTCTTCAATTTTACAGCTGA 2644
|||||
```

Search completed: January 6, 2004, 16:13:03  
Job time : 1111 secs

```
Db 283 CAAAAACACATTTCTAAATAGTAGTAACTTTTCTTAAATTTCTTCAATTTTACAGCTGA 224
QY 2645 TGATATGCATAAGTGCTGTAAAGGCTTGTAACTGGGAAATATTCCACCTGATAATAGCCC 2704
Db 223 TGATATGCATAAGTGCTGTAAAGGCTTGTAACTGGGAAATATTCCACTGATAATAGCCC 164
QY 2705 AGATTCTACTGTATTCCCAAAAGGCAATATTAAAGGTAGATAGATGATTAGTAGTATTG 2764
Db 163 AGATTCTACTGTATCCCCCAAAAGGCAATATTAAAGGTAGATAGATGATTAGTAGTATTG 104
QY 2765 TTACACACTATTTTGGAAATTAGAGAAACATACAGAAAGAAATTTAGGGGCTTAAACATTACG 2824
Db 103 TTACACACTATTTTGGAAATTAGAGAAACATACAGAAAGAAATTTAGGGGCTTAAACATTGCG 44
QY 2825 ACTGAATGCACCTTTTAGTATATAAGGGGCACAGTTTGTATATTTT 2867
Db 43 ACTGAATGCACCTTTTAGTATATAAGGGGCACAGTTTGTATATTTT 1
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: January 6, 2004, 10:12:26 ; Search time 903 Seconds  
(without alignments)  
11075.767 Million cell updates/sec

Title: US-09-890-549-16  
Perfect score: 3705  
Sequence: 1 ggcgcaggagcagtagtg.....atttgaaaaaaaaaaaaa 3705

Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_19Jun03:.\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3705	100.0	3705	21	Sequence encoding
2	3611	97.5	3664	23	Human transcriptio
3	3412	92.1	3711	21	Human OREF3698
4	2596	70.1	2783	20	Human breast tumou
5	2589	69.9	2749	24	Human cDNA encodin
6	2565	69.2	2905	22	Human polynucleoti
7	2130	57.5	3650	22	Human EPLIN (epith
8	2092	56.5	2267	21	Human secreted pro

9	2041	55.1	3543	22	AAF555696	Human EPLIN (epith
10	1519	41.0	2207	22	AAH18125	Human cDNA sequenc
11	1310	35.4	1713	22	AAI59955	Human polynucleoti
12	1296	35.0	1754	24	ABI98973	Human cancer suppr
13	1269	34.3	2158	21	AAI59973	Human secreted pro
14	1186	32.0	1567	24	AAI58588	Human secreted pro
15	604	16.3	641	20	AAZ52892	Human encoding huma
16	590	15.9	3465	23	AAZ52892	Human prostate tum
17	556	15.0	732	21	AAZ52892	DNA encoding novel
18	466	12.6	1014	23	AAZ52892	Human pancreatic c
19	415	11.2	565	24	AAZ52892	DNA encoding novel
20	385	10.4	698	22	AAH72864	Human pancreatic c
21	362	9.8	586	21	AAH72864	Human cervical can
22	343	9.3	343	24	ABK29495	Human colon cancer
23	343	9.3	343	24	ABK29495	Human colon cancer
24	340	9.2	463	24	ABQ58707	Colon adenocarcino
25	335	8.8	389	21	AAZ52892	Human colon cancer
26	314	8.5	376	21	AAZ52892	Human secreted pro
27	276	7.4	279	25	AAZ52892	DNA encoding novel
28	248	6.7	413	25	AAZ52892	Chinese hamster si
29	236	6.4	532	22	AAH12979	Chinese hamster si
30	216	6.2	419	22	AAH12979	Human cDNA clone (
31	210	5.8	408	23	AAZ52892	Human Ca2+ pump PM
32	212	5.7	247	22	AAH69368	DNA encoding novel
33	209	5.6	231	22	AAH72234	Human cervical can
34	209	5.6	255	22	AAH72234	Human cervical can
35	208	5.6	265	22	AAZ52892	Human ovarian canc
36	186	5.0	283	24	ABV88456	Human colon cancer
37	175	4.7	178	24	ABL38497	Human colon tumour
38	164	4.4	2226	23	AAZ52892	DNA encoding novel
39	158	4.3	382	24	ABQ58853	Human colon cancer
40	155	4.2	232	22	AAH19758	Human breast cance
41	143	3.9	296	21	AAZ52892	Human secreted pro
42	140	3.8	269	22	AAH10178	Human breast cance
43	136	3.7	996	23	AAZ52892	DNA encoding novel
44	129	3.5	577	23	ABV49283	Human prostate exp
45	119	3.2	238	25	AAZ52892	Chinese hamster si

ALIGNMENTS

RESULT 1  
AAZ53826  
ID AAZ53826 standard; DNA; 3705 BP.  
XX AAZ53826;  
AC AAZ53826;  
XX AAZ53826;  
DT 03-JAN-2001 (first entry)  
XX Sequence encoding lipid associated protein (LIPAP) 2766980CB1.  
DE Lipid associated protein; LIPAP; treatment; prophylaxis; agonist;  
XX antagonist; antibody; cardiovascular disease; neurological disease;  
KW gastrointestinal disease; lipid metabolism; detection;  
KW amplification; monitoring; hybridisation; antisense; triplex;  
KW ribozyme; screening; immunoassay; ds.  
XX Homo sapiens.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FT CDS 137..2416  
FT /\*tag= a  
FT /product= Lipid associated protein  
XX WO200049043-A2.  
PN WO200049043-A2.  
XX 24-AUG-2000.  
PD 24-AUG-2000.  
XX 18-FEB-2000; 2000WO-US04160.  
PP 18-FEB-2000; 2000WO-US04160.  
XX 19-FEB-1999; 99US-0120703.  
PR 08-JUL-1999; 99US-0142762.

XX (INCY-) INCYTE PHARM INC.  
XX Tang YT, Hillman JL, Yue H, Azimzai Y, Baughn MR, Tran B;  
XX WPI; 2000-549264/50.  
DR P-PSDB; AAY97286.  
XX  
PT New human lipid-associated proteins, nucleic acids, and antibodies,  
PT useful for diagnosis, treatment and prevention of e.g. cardiovascular  
PT disease  
XX  
XX Claim 4; Page 87-88; 93pp; English.  
XX  
CC Lipid-associated proteins (LIPAP) can be used for treating or  
CC preventing disorders associated with decreased expression of LIPAP,  
CC for screening for agonists or antagonists of LIPAP, and to raise  
CC specific antibodies. Antagonists and antagonists of LIPAP are useful  
CC for treating diseases associated with reduced or increased levels of  
CC LIPAP, e.g. cardiovascular, neurological and gastrointestinal  
CC diseases and disorders of lipid metabolism. Fragments of the nucleic  
CC acid encoding LIPAP are useful for detection of full length coding  
CC sequences, in hybridization and/or amplification assays or for  
CC diagnosis or monitoring. Nucleotides encoding LIPAP are used  
CC to screen for compounds that specifically modify LIPAP expression,  
CC for recombinant production of LIPAP, in gene therapy, as a source of  
CC therapeutic antisense, triplex-forming, or ribozyme agents and for  
CC genomic mapping. Antibodies to the proteins are used for diagnosis  
CC and monitoring of LIPAP-associated disease by immunoassay, as  
CC antagonists, in competitive drug screens and for affinity  
CC purification of natural LIPAP.  
XX  
SQ Sequence 3705 BP; 1217 A; 743 C; 826 G; 919 T; 0 other;

Query Match  
Best Local Similarity 100.0%; Score 3705; DB 21; Length 3705;  
Matches 3705; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGCGCAGGAGCAGTAGGTTAGCAGCTTGGTCGCGACAGTGCGCTAGGTAGAGCGCC 60  
Db |||||  
QY 61 GGGACCTGTGACAGGCGCTGGTAGCAGCGCAGAGGAAAGCGGCTTTTAGCCAGGTATTTC 120  
Db |||||  
QY 61 GGGACCTGTGACAGGCGCTGGTAGCAGCGCAGAGGAAAGCGGCTTTTAGCCAGGTATTTC 120  
Db |||||  
QY 121 AGTGCTGTGACAGAGATGGAAATCATCTCCATTTAATAGAGCGGAATGGACCTCACTATC 180  
Db |||||  
QY 181 ATTGAGGCTAACAGCAAGAACTTTCTTGTGTCAACAGAAAGTCAATCGGCTATTGT 240  
Db |||||  
QY 181 ATTGAGGCTAACAGCAAGAACTTTCTTGTGTCAACAGAAAGTCAATCGGCTATTGT 240  
Db |||||  
QY 241 GGAATATTCTCCAAAGTACCAGAGCAGCTGAGCAACAAACATGGAGAGAGAGAAAG 300  
Db |||||  
QY 241 GGAATATTCTCCAAAGTACCAGAGCAGCTGAGCAACAAACATGGAGAGAGAGAAAG 300  
Db |||||  
QY 301 TAAACCCGAAATCTCTCCAGCAGCTTTAGAAAGGGACCCCTGACTGTGTAAAGAGAA 360  
Db |||||  
QY 301 TAAACCCGAAATCTCTCCAGCAGCTTTAGAAAGGGACCCCTGACTGTGTAAAGAGAA 360  
Db |||||  
QY 361 GTGGAGAACCCAGGCTGGGAGCAGAGTCTCACAGAGCTCTCTACGGAAACAGCAGAC 420  
Db |||||  
QY 421 TGAGATTAGGCACAGACAGACCACTCTCTCTGTGTGAAGTGAACGCCACGCTCTTGG 480  
Db |||||  
QY 421 TGAGATTAGGCACAGACAGACCACTCTCTCTGTGTGAAGTGAACGCCACGCTCTTGG 480  
Db |||||  
QY 481 AGCCAAAGCTGACCAAGAAAGAAATCCACCCCGAGATCTAGACTCAGGTCACCTCTGA 540  
Db |||||  
QY 481 AGCCAAAGCTGACCAAGAAAGAAATCCACCCCGAGATCTAGACTCAGGTCACCTCTGA 540  
Db |||||

QY 541 AGCCCTCGTTGAGGTCGATATCCCCACATCAAGACGGTGAAGGATCTTAAAGACCACTC 600  
Db |||||  
QY 541 AGCCCTCGTTGAGGTCGATATCCCCACATCAAGACGGTGAAGGATCTTAAAGACCACTC 600  
Db |||||  
QY 601 AACAGAAAGTAAAGAAATGGAAATTTGCTAGGAGAAATCCAGGCATGAAGTAGAAAAATC 660  
Db |||||  
QY 601 AACAGAAAGTAAAGAAATGGAAATTTGCTAGGAGAAATCCAGGCATGAAGTAGAAAAATC 660  
Db |||||  
QY 661 AGAAATCAGTGAAGAACACAGATGCTTCGGGCAAAATAGAGAAATATATATGTTCCGCTGAA 720  
Db |||||  
QY 661 AGAAATCAGTGAAGAACACAGATGCTTCGGGCAAAATAGAGAAATATATATGTTCCGCTGAA 720  
Db |||||  
QY 721 CAGGCTTAAGATGATGTTTGAGAAAGGTGAACCAACTCAAACTAAGATTTCTCCGGGCCCA 780  
Db |||||  
QY 721 CAGGCTTAAGATGATGTTTGAGAAAGGTGAACCAACTCAAACTAAGATTTCTCCGGGCCCA 780  
Db |||||  
QY 781 AAGCGAAGTCAAGTGAAGAAAGATCTCTGAAAAACAGCTATTCTCTAGATGACCTGGA 840  
Db |||||  
QY 781 AAGCGAAGTCAAGTGAAGAAAGATCTCTGAAAAACAGCTATTCTCTAGATGACCTGGA 840  
Db |||||  
QY 841 AATAGGCCAGGTCAGTTGTCATCTTCTAGATTTGACTCGGAGAAAAATGAGAGTAGACG 900  
Db |||||  
QY 841 AATAGGCCAGGTCAGTTGTCATCTTCTAGATTTGACTCGGAGAAAAATGAGAGTAGACG 900  
Db |||||  
QY 901 AAATCTGGAATCTCCACGCTCTCAGAAACCTCTATTAAGATCGAATGCGCAAGTACCA 960  
Db |||||  
QY 901 AAATCTGGAATCTCCACGCTCTCAGAAACCTCTATTAAGATCGAATGCGCAAGTACCA 960  
Db |||||  
QY 961 GGCAGCTGTGTCACAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1020  
Db |||||  
QY 961 GGCAGCTGTGTCACAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1020  
Db |||||  
QY 1021 TGGCGAAATCAAAATTCATTAATATGGAGCAAAAGGAAATGTGCCCCAGGTCTCAGGT 1080  
Db |||||  
QY 1021 TGGCGAAATCAAAATTCATTAATATGGAGCAAAAGGAAATGTGCCCCAGGTCTCAGGT 1080  
Db |||||  
QY 1081 CTGCATCACCCTCAGAGAGGGGAAAGATTTCTGCAATAGAAATAGCCTGGCAGTCCG 1140  
Db |||||  
QY 1081 CTGCATCACCCTCAGAGAGGGGAAAGATTTCTGCAATAGAAATAGCCTGGCAGTCCG 1140  
Db |||||  
QY 1141 TTCACCCCTGCCGAAAGATGACTCCCGTGACTCCAGGTTAAGAGTGAAGTTCGCAAGTCCG 1200  
Db |||||  
QY 1141 TTCACCCCTGCCGAAAGATGACTCCCGTGACTCCAGGTTAAGAGTGAAGTTCGCAAGTCCG 1200  
Db |||||  
QY 1201 TGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGAGCTTCTTCTGAAAGTCTCC 1260  
Db |||||  
QY 1201 TGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGAGCTTCTTCTGAAAGTCTCC 1260  
Db |||||  
QY 1261 TCCGAAAGCAATGAAGAAAGTTTCAGGCACCTGCAAGAGAGACCTGCGTGGAAATGTGAGAA 1320  
Db |||||  
QY 1261 TCCGAAAGCAATGAAGAAAGTTTCAGGCACCTGCAAGAGAGACCTGCGTGGAAATGTGAGAA 1320  
Db |||||  
QY 1321 GACAGTCTATCCAAATGGAGCGTCTCTTGGCCCAACAGCAGGTTGTTTCACATCAGTGTCTT 1380  
Db |||||  
QY 1321 GACAGTCTATCCAAATGGAGCGTCTCTTGGCCCAACAGCAGGTTGTTTCACATCAGTGTCTT 1380  
Db |||||  
QY 1381 CGTTGCTCTTATGGCAACAACAACTCAGTCTAGAACATATGCATCTTTACATGGAAG 1440  
Db |||||  
QY 1381 CGTTGCTCTTATGGCAACAACAACTCAGTCTAGAACATATGCATCTTTACATGGAAG 1440  
Db |||||  
QY 1441 AATCTATTAGGCTCACTTCAATCAACTCTTTAAATCTTAAGGGCAACTATGATGAAGG 1500  
Db |||||  
QY 1441 AATCTATTAGGCTCACTTCAATCAACTCTTTAAATCTTAAGGGCAACTATGATGAAGG 1500  
Db |||||  
QY 1501 CTTTGGGCACAGCACACAGGATCTATGGCAAGCAAAATGAAAAACGAGAGATTTT 1560  
Db |||||  
QY 1501 CTTTGGGCACAGCACACAGGATCTATGGCAAGCAAAATGAAAAACGAGAGATTTT 1560  
Db |||||  
QY 1561 GGAGAGACAGGCCAGCTTGCAAAATGCAAGGAGACCCCTCACAGCCCGGGGTAGAAGA 1620  
Db |||||  
QY 1561 GGAGAGACAGGCCAGCTTGCAAAATGCAAGGAGACCCCTCACAGCCCGGGGTAGAAGA 1620  
Db |||||



```
QY 1621 TGCCCTCTATTGCTAAGGTGGGTGCTCGCTGCAAGTATGGAAGCCAGGCGCTCTCTCA 1680
Db 1621 TGCCCTCTATTGCTAAGGTGGGTGCTCGCTGCAAGTATGGAAGCCAGGCGCTCTCTCA 1680
QY 1681 GCAGGAGGAAGGAACAAGCCAGCTGAAACCAAGAAAGCTGAGGATCGCCTCGCCACCCCCC 1740
Db 1681 GCAGGAGGAAGGAACAAGCCAGCTGAAACCAAGAAAGCTGAGGATCGCCTCGCCACCCCCC 1740
QY 1741 CACTGAACCTTGGAGTTTCAGGAAGTGCCTTGGAGGAAGGGATCAAAATGTCAAAAGCCCAA 1800
Db 1741 CACTGAACCTTGGAGTTTCAGGAAGTGCCTTGGAGGAAGGGATCAAAATGTCAAAAGCCCAA 1800
QY 1801 ATGGCTCTCTGAAGACGAATCAGCAAGCCCGAAGTTCTCGAGGATGCGATCTAGATCT 1860
Db 1801 ATGGCTCTCTGAAGACGAATCAGCAAGCCCGAAGTTCTCGAGGATGCGATCTAGATCT 1860
QY 1861 GAAGAAGCTAAGACCATCTCTTCACTGAAGGAAGAGCCGCCCATTCAGTGTAGCAGC 1920
Db 1861 GAAGAAGCTAAGACCATCTCTTCACTGAAGGAAGAGCCGCCCATTCAGTGTAGCAGC 1920
QY 1921 TTCAATTTCAAAGCACCTCTGTCAAGAGCCCAAAACCTGTGTCCCACTCTATCAGGAAGG 1980
Db 1921 TTCAATTTCAAAGCACCTCTGTCAAGAGCCCAAAACCTGTGTCCCACTCTATCAGGAAGG 1980
QY 1981 CTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGGTGGAAGTTGCAAGAAAGGAACA 2040
Db 1981 CTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGGTGGAAGTTGCAAGAAAGGAACA 2040
QY 2041 AGTGGAAATTCGCAAGGCTTCTAGAGAAGATGGAATGTGGGAAACCAACCTGGCAAAA 2100
Db 2041 AGTGGAAATTCGCAAGGCTTCTAGAGAAGATGGAATGTGGGAAACCAACCTGGCAAAA 2100
QY 2101 CAAAGAAATCTAAAGGAGAGACAGGAAGAGAGTAAGGAAGGTCTAGTTTGGAGATGGA 2160
Db 2101 CAAAGAAATCTAAAGGAGAGACAGGAAGAGAGTAAGGAAGGTCTAGTTTGGAGATGGA 2160
QY 2161 GAATGAGAAATCTTGTAAGAAATGGTGAGACTCCGATGAAGATGATAACAGCTTCTCTAA 2220
Db 2161 GAATGAGAAATCTTGTAAGAAATGGTGAGACTCCGATGAAGATGATAACAGCTTCTCTAA 2220
QY 2221 ACAACAAATCTCCACAGAACCCCAAGTCTCTGAATTTGGTGGAGTTTGTAGACAAACCTT 2280
Db 2221 ACAACAAATCTCCACAGAACCCCAAGTCTCTGAATTTGGTGGAGTTTGTAGACAAACCTT 2280
QY 2281 TGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCAGAGATGTGGAATCTGGAGGGAGA 2340
Db 2281 TGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCAGAGATGTGGAATCTGGAGGGAGA 2340
QY 2341 AGTGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATCAGGA 2400
Db 2341 AGTGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAATCGGTATTATGATCAGGA 2400
QY 2401 TGAGGATGAAGAGTGACAAATTTGCAATGATGCTGGGCCCTTAAATTCATGTTAGTTAGC 2460
Db 2401 TGAGGATGAAGAGTGACAAATTTGCAATGATGCTGGGCCCTTAAATTCATGTTAGTTAGC 2460
QY 2461 GAGCCACTCCCTTTGTCAAAATGTGATGCACATGAAGAGGTATCCAGCATGAAATGTA 2520
Db 2461 GAGCCACTCCCTTTGTCAAAATGTGATGCACATGAAGAGGTATCCAGCATGAAATGTA 2520
QY 2521 ATTTACTTGGAAAGTAACCTTTGAAAAGAAATTCCTTCTTAAATTCAAAATCAAAACAAAAA 2580
Db 2521 ATTTACTTGGAAAGTAACCTTTGAAAAGAAATTCCTTCTTAAATTCAAAATCAAAACAAAAA 2580
QY 2581 AACACAAAAACACATTTCTAAATCTAGAGATAACTTTACTTTAAATTCCTTCAATTTAGCA 2640
Db 2581 AACACAAAAACACATTTCTAAATCTAGAGATAACTTTACTTTAAATTCCTTCAATTTAGCA 2640
QY 2641 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
Db 2641 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
QY 2701 GCCCAGATTCTACTGTATTCCCAAAAGGCAATATTAAAGGTAGATAGATTAGTAGTAT 2760
```

```
Db 2701 GCCCAGATTCTACTGTATTCCCAAAAGGCAATATTAAAGGTAGATAGATTAGTAGTAT 2760
QY 2761 ATTGTTACACACTATTTTTGAATTTAGAGAACATACAGAAAGAAATTTAGGGGCTTTAAACAT 2820
Db 2761 ATTGTTACACACTATTTTTGAATTTAGAGAACATACAGAAAGAAATTTAGGGGCTTTAAACAT 2820
QY 2821 TAGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATATTTTAAATGAATACCAA 2880
Db 2821 TAGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATATTTTAAATGAATACCAA 2880
QY 2881 TTTAAATTTTGTAGTATTACCTGTTAAAGAGATTTTAGTCTTTAAATTTTGTAGTTAA 2940
Db 2881 TTTAAATTTTGTAGTATTACCTGTTAAAGAGATTTTAGTCTTTAAATTTTGTAGTTAA 2940
QY 2941 TTTTCTTGCTGTGATATATATGAGGAATTTACTACTTTTATGCTGCTCTCTAAACTACA 3000
Db 2941 TTTTCTTGCTGTGATATATATGAGGAATTTACTACTTTTATGCTGCTCTCTAAACTACA 3000
QY 3001 TCTGAACTCGACGCTCTCGAGGTATAATAACAAGAGACACTTTTGGGCAATTTGAAAAA 3060
Db 3001 TCTGAACTCGACGCTCTCGAGGTATAATAACAAGAGACACTTTTGGGCAATTTGAAAAA 3060
QY 3061 CCAACCTACACTCTTCCGCTGTAGAGAGATCTGCTGCCCAATTAAGCTTTTGTATC 3120
Db 3061 CCAACCTACACTCTTCCGCTGTAGAGAGATCTGCTGCCCAATTAAGCTTTTGTATC 3120
QY 3121 TGCCAGTGAATTTACTGTACTCCAAATGATTCCTTTCTTTCTGTGTATCTGTGCTTC 3180
Db 3121 TGCCAGTGAATTTACTGTACTCCAAATGATTCCTTTCTTTCTGTGTATCTGTGCTTC 3180
QY 3181 TCATAAATCTGAAAGCTGCAATATTTTAGTAATACCTTCGGGATCAGTGTCCCCATCT 3240
Db 3181 TCATAAATCTGAAAGCTGCAATATTTTAGTAATACCTTCGGGATCAGTGTCCCCATCT 3240
QY 3241 TCCGTGTTAGAGCAAGTGAAGAGTTTAAAGGAGGAAGAAAGAACTGTCTTACACA 3300
Db 3241 TCCGTGTTAGAGCAAGTGAAGAGTTTAAAGGAGGAAGAAAGAACTGTCTTACACA 3300
QY 3301 CTGGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGATGTCCCTTTTGGAGACACTA 3360
Db 3301 CTGGAGCTCAGACCTCTAAACCCCTGATTTCCCTTATGATGTCCCTTTTGGAGACACTA 3360
QY 3361 ATTTTAAATCTTACTAGCTCTGAAATATATTTGATTTTATCAGATATTTCTCAGGGTG 3420
Db 3361 ATTTTAAATCTTACTAGCTCTGAAATATATTTGATTTTATCAGATATTTCTCAGGGTG 3420
QY 3421 AAATTAACCAACTATAGGCCCTTTTCTGGGATGATTTTCTAGTCTTAAGGTTTGGGA 3480
Db 3421 AAATTAACCAACTATAGGCCCTTTTCTGGGATGATTTTCTAGTCTTAAGGTTTGGGA 3480
QY 3481 CATTATAAATCTTGAGTACATTTGTTGACACAGTTGATATTTCCAAATTTGATGGATGGA 3540
Db 3481 CATTATAAATCTTGAGTACATTTGTTGACACAGTTGATATTTCCAAATTTGATGGATGGA 3540
QY 3541 GGGAGAGGTCTCTTAAGCTGTAGGCTTTTCTTGTACTGCAATTTATAGAGATTTAGCTTT 3600
Db 3541 GGGAGAGGTCTCTTAAGCTGTAGGCTTTTCTTGTACTGCAATTTATAGAGATTTAGCTTT 3600
QY 3601 AATATTTTTTAGAGATGATAAAACATCTGCTTTCTTTAGTCTTACCTAGTCTGAAACATTT 3660
Db 3601 AATATTTTTTAGAGATGATAAAACATCTGCTTTCTTTAGTCTTACCTAGTCTGAAACATTT 3660
QY 3661 TTATTCATTAAGATTTTAAATTTAAATTTGAAAAAATTTGAAAAAATTTGAAAAAATTT 3705
Db 3661 TTATTCATTAAGATTTTAAATTTAAATTTGAAAAAATTTGAAAAAATTTGAAAAAATTT 3705
```

RESULT 2  
ABX71307  
ID ABX71307 standard; cDNA; 3664 bp.  
XX  
AC ABX71307;

XX  
 DT 14-APR-2003 (first entry)  
 XX Human transcription factor cDNA from clone DKFZphutcl\_18c19.  
 DE Human; gene; Gene therapy; vaccine; disease treatment; detection; ss.  
 XX Homo sapiens.  
 OS  
 XX W0200112659-A2.  
 PN 22-FEB-2001.  
 XX  
 PD 18-AUG-2000; 2000WO-IB01496.  
 XX  
 PF 18-AUG-1999; 99US-0149499.  
 XX 28-SEP-1999; 99US-0156503.  
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.  
 PA  
 XX  
 XX  
 PI Wiemann S;  
 XX  
 XX WPI; 2001-327840/34.  
 DR P-PSDB; ABU52869.  
 XX  
 XX Nucleic acids having the sequences of clones isolated from libraries of  
 PT different human tissues, useful in recombinant DNA methodologies -  
 PT  
 XX Claim 1; Page 449-450; 1095pp; English.  
 PS  
 XX This invention describes novel polynucleotides and polypeptides isolated  
 CC from human cDNA libraries which can be used for gene therapy or in  
 CC vaccines. The polynucleotides of the invention and antibodies encoded by  
 CC them may be used in the prevention, diagnosis and treatment of diseases  
 CC associated with inappropriate polypeptide expression. The products of the  
 CC invention may also be used to identify modulators of expression and  
 CC activity and to down regulate expression and activity. The antibodies of  
 CC the invention may also be used as diagnostic agents for detecting the  
 CC presence of polypeptides in samples. This sequence encodes a polypeptide  
 CC described in the disclosure of the invention.  
 XX  
 XX Sequence 3664 BP; 1210 A; 734 C; 809 G; 911 T; 0 other;

Query Match 97.5%; Score 3611; DB 23; Length 3664;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3661; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 GCGCTAGGTAGCGCGCGGACCTGTGACAGGGCTGGTAGCGGCAGAGGAAGCGGC 103  
 DB 1 GCGCTAGGTAGAGCGCGCGGACCTGTGACAGGGCTGGTAGCGGCAGAGGAAGCGGC 60  
 QY 104 TTTTAGCCAGGTATTTTCAGTGTCTGTAGACAGATGGAATCATCTCCATTTAATAGCGG 163  
 DB 61 TTTTAGCCAGGTATTTTCAGTGTCTGTAGACAGATGGAATCATCTCCATTTAATAGCGG 120  
 QY 164 CAATGSGACCTCACTATCAATGAGGGTAACAGCCAAAGAACTTCTTCTGTGCAACAAGAC 223  
 DB 121 CAATGSGACCTCACTATCAATGAGGGTAACAGCCAAAGAACTTCTTCTGTGCAACAAGAC 180  
 QY 224 AAGTCATCGGCTATTGTGGAATATTTCTCCAAAGTACAGAAAGCAGCTGAAGAAACAAAC 283  
 DB 181 AAGTCATCGGCTATTGTGGAATATTTCTCCAAAGTACAGAAAGCAGCTGAAGAAACAAAC 240  
 QY 284 ATGGAGAAGAGAGAGTACACCGGAGAACCCAGGGCTGGGACGAGTCTCACACAGCTCT 343  
 DB 241 ATGGAGAAGAGAGAGTACACCGGAGAACCCAGGGCTGGGACGAGTCTCACACAGCTCT 300  
 QY 344 ACTGTGTTAAAGAGAGTGGGAGAACCCAGGGCTGGGACGAGTCTCACACAGCTCT 403  
 DB 301 ACTGTGTTAAAGAGAGTGGGAGAACCCAGGGCTGGGACGAGTCTCACACAGCTCT 360  
 QY 404 CTACGGAACAGCAGCACTGAGATTAGGCAACAGAGCAATCCTCTCTGAGTGACA 463  
 DB

DB 361 CTACGGAACAGCAGCACTGAGATTAGGCAACAGACCACTCTCTCTGCTGAAAGTGACA 420  
 QY 464 AGCCACGCTGTTCTGGAGCCAAAGCTGACCAAGAGAAACAAATCCACCTCCAGATCTAGA 523  
 DB 421 AGCCACGCTGTTCTGGAGCCAAAGCTGACCAAGAGAAACAAATCCACCTCCAGATCTAGA 480  
 QY 524 CTCAGGTACCTCTCTGAAAGCCCTGTTAGGGTCGATATCCCCACATCAAGGACGGTGAG 583  
 DB 481 CTCAGGTACCTCTCTGAAAGCCCTGTTAGGGTCGATATCCCCACATCAAGGACGGTGAG 540  
 QY 584 GATCTTAAAGACCACTCAACAGAGAAAGTAAAGAAATGAAAAATTTGCTTAGAGAAATCCAGG 643  
 DB 541 GATCTTAAAGACCACTCAACAGAGAAAGTAAAGAAATGAAAAATTTGCTTAGAGAAATCCAGG 600  
 QY 644 CATGAAGTAGAAAAATCAGAAATCAGTGAAGAAACACAGATGCTTCGGGCAAAATAGAGAAA 703  
 DB 601 CATGAAGTAGAAAAATCAGAAATCAGTGAAGAAACACAGATGCTTCGGGCAAAATAGAGAAA 660  
 QY 704 TATAATGTTCCGCTGAAACAGGCTTAAGATGATGTTTGAAGAAAGTGAACCAATCAAACT 763  
 DB 661 TATAATGTTCCGCTGAAACAGGCTTAAGATGATGTTTGAAGAAAGTGAACCAATCAAACT 720  
 QY 764 AAGATTCCTCGGGCCCAAGCCGAAAGTGAAGTGAAGGAAAGATCTCTGAAAAACAGCTAT 823  
 DB 721 AAGATTCCTCGGGCCCAAGCCGAAAGTGAAGTGAAGGAAAGATCTCTGAAAAACAGCTAT 780  
 QY 824 TCTCTAGATCACTTGGAAATAGCCCGAGGTGCTGTCATCTTCTACATTTGACTCGGAG 893  
 DB 781 TCTCTAGATCACTTGGAAATAGCCCGAGGTGCTGTCATCTTCTACATTTGACTCGGAG 840  
 QY 884 AAAAAATGAGAGTAGAGCAAAATCTGGAACCTTCACAGGCTCTCAGAAACCTCTATTAAGGAT 943  
 DB 841 AAAAAATGAGAGTAGAGCAAAATCTGGAACCTTCACAGGCTCTCAGAAACCTCTATTAAGGAT 900  
 QY 944 CGAATGGCCAAAGTACAGGAGCTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAAT 1003  
 DB 901 CGAATGGCCAAAGTACAGGAGCTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAAT 960  
 QY 1004 GAGCTCAAAAGCCAGTGTGTCGCAAAATCAAAATTCATAAAATGGAGCAAAAGAGAAATGTG 1063  
 DB 961 GAGCTCAAAAGCCAGTGTGTCGCAAAATCAAAATTCATAAAATGGAGCAAAAGAGAAATGTG 1020  
 QY 1064 CCCCCAGGTCCTGAGGTCTGCAATCAACCCATCAGGAAGGGGAAAGATTTCTCCAAATGAG 1123  
 DB 1021 CCCCCAGGTCCTGAGGTCTGCAATCAACCCATCAGGAAGGGGAAAGATTTCTCCAAATGAG 1080  
 QY 1124 AATAGCCTGGCAGTCTGTTCCACCCCTCCCGAAGATGACTCCCGTGACTCCCGAGGTTAAG 1183  
 DB 1081 AATAGCCTGGCAGTCTGTTCCACCCCTCCCGAAGATGACTCCCGTGACTCCCGAGGTTAAG 1140  
 QY 1184 AGTGAGGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGT 1243  
 DB 1141 AGTGAGGTTCAACAGCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGT 1200  
 QY 1244 CTTTCTGAAAGTTCTCCTCCCAAGCAATGAAGAGTTTCAGGCACCTGCAAGAGAGACC 1303  
 DB 1201 CTTTCTGAAAGTTCTCCTCCCAAGCAATGAAGAGTTTCAGGCACCTGCAAGAGAGACC 1260  
 QY 1304 TCGCTGGAATGTGAGAAGACAGTCTATCCAATGGAGCGTCTCTGGCCCAACAGCAGGTG 1363  
 DB 1261 TCGCTGGAATGTGAGAAGACAGTCTATCCAATGGAGCGTCTCTGGCCCAACAGCAGGTG 1320  
 QY 1364 TTTTCAATCAGCTGCTCGTTCCTTCTTATGCAACCAACAACTCAGTCTAGGAAACATAT 1423  
 DB 1321 TTTTCAATCAGCTGCTCGTTCCTTCTTATGCAACCAACAACTCAGTCTAGGAAACATAT 1380  
 QY 1424 GCATCTTTTACATGAAGAAATCTATTGTAAGCCTCACTTTCAATCAACTCTTTTAAATCTAAG 1483  
 DB 1391 GCATCTTTTACATGAAGAAATCTATTGTAAGCCTCACTTTCAATCAACTCTTTTAAATCTAAG 1440  
 QY 1484 GGCAACTATGATGAAGGCTTTGGGCACAGACCAACAGAGATCTATGGGCAAGCAAAAAAT 1543  
 DB 1441 GGCAACTATGATGAAGGCTTTGGGCACAGACCAACAGAGATCTATGGGCAAGCAAAAAAT 1500

QY 1544 GAAACGAAAGAGATTTTGGAGAGACAGCCCGAGCTTGCATAATGCAAGGGAGAGACCCCTCAC 1603  
DB 1501 GAAACGAAAGAGATTTTGGAGAGACAGCCCGAGCTTGCATAATGCAAGGGAGAGACCCCTCAC 1560  
QY 1604 AGCCGAGGGGTAGAGATGCCCTATTGCTAGAGTGGGTGCTTGGCTGCAAGTAGTAA 1663  
DB 1561 AGCCGAGGGGTAGAGATGCCCTATTGCTAGAGTGGGTGCTTGGCTGCAAGTAGTAA 1620  
QY 1664 GCAAGGCCCTCTCTCAGCAGGAGAGAGAAAGCAAGCCAGCTGAAACCAAGAGAGCTGAGG 1723  
DB 1621 GCAAGGCCCTCTCTCAGCAGGAGAGAGAGAAAGCAAGCCAGCTGAAACCAAGAGAGCTGAGG 1680  
QY 1724 ATCGGCTGGCCACCCGCCACTGAACTTGGAAAGTTCAGGAAGTGCCTTGGAGGAAGGGATC 1783  
DB 1681 ATCGGCTGGCCACCCGCCACTGAACTTGGAAAGTTCAGGAAGTGCCTTGGAGGAAGGGATC 1740  
QY 1784 AAAATGTCAAGCCCAATGGCCCTCTCAAGAGCGAAATCAGCAAGCCCGAAGTTCCTGAG 1843  
DB 1741 AAAATGTCAAGCCCAATGGCCCTCTCAAGAGCGAAATCAGCAAGCCCGAAGTTCCTGAG 1800  
QY 1844 GATGTCGATCTAGATCTCAAGAGCTAAGACGATCTTCAAGAGCCCAAGAACTGTGTCC 1903  
DB 1801 GATGTCGATCTAGATCTCAAGAGCTAAGACGATCTTCAAGAGCCCAAGAACTGTGTCC 1860  
QY 1904 CCATTCACTGTAGCAGCTTCAATTTCAAGACACCTCTGTCAAGAGCCCAAGAACTGTGTCC 1963  
DB 1861 CCATTCACTGTAGCAGCTTCAATTTCAAGACACCTCTGTCAAGAGCCCAAGAACTGTGTCC 1920  
QY 1964 CCACCTATCAGGAAGGCTGGAGCATGTCAAGAGAGTGAAGAGTCTGTGGGTGGAAGA 2023  
DB 1921 CCACCTATCAGGAAGGCTGGAGCATGTCAAGAGAGTGAAGAGTCTGTGGGTGGAAGA 1980  
QY 2024 GTTGAGAAAGGAACAAAGTGGAAATCTTAAGAGAGACAGGAGAGAGTGAAGAGT 2083  
DB 1981 GTTGAGAAAGGAACAAAGTGGAAATCTTAAGAGAGTGAAGAGTGAAGAGT 2040  
QY 2084 AAAACAACCTGGCAAAACAAAGATCTTAAGAGAGACAGGAGAGAGTGAAGAGT 2143  
DB 2041 AAAACAACCTGGCAAAACAAAGATCTTAAGAGAGACAGGAGAGAGTGAAGAGT 2100  
QY 2144 CATAGTTGGAGATGGAGATAGAAATCTTGAAGAAATGGTGAGACTCCGATGAAGAT 2203  
DB 2101 CATAGTTGGAGATGGAGATAGAAATCTTGAAGAAATGGTGAGACTCCGATGAAGAT 2160  
QY 2204 GATAACAGCTTCTCAACACAACTCTCCACAGAACCCCAAGTCTCTGAATTCGTCAGT 2263  
DB 2161 GATAACAGCTTCTCAACACAACTCTCCACAGAACCCCAAGTCTCTGAATTCGTCAGT 2220  
QY 2264 TTTGTAGACAACACCTTTGCTGAAGAAATCACTACTCAGAAATCCCAAGTCTCTGAATTCGTCAGT 2323  
DB 2221 TTTGTAGACAACACCTTTGCTGAAGAAATCACTACTCAGAAATCCCAAGTCTCTGAATTCGTCAGT 2280  
QY 2324 GAACTCTGGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATATAAGAGAAAT 2383  
DB 2281 GAACTCTGGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATATAAGAGAAAT 2340  
QY 2384 CGGTATTATGATGAGGATGAGGATGAAGAGTGCAAAATTCGAATGATGCTGGGCCCTTAAA 2443  
DB 2341 CGGTATTATGATGAGGATGAGGATGAAGAGTGCAAAATTCGAATGATGCTGGGCCCTTAAA 2400  
QY 2444 TTTATGTTAGTGTAGCGGCCACTGCCCTTTGTCAAATGTGATGCAATTAAGCAGGTA 2503  
DB 2401 TTTATGTTAGTGTAGCGGCCACTGCCCTTTGTCAAATGTGATGCAATTAAGCAGGTA 2460  
QY 2504 TCCCAGATGAAATGTAATTTACTTGGAAAGTAACTTTGGAAGAAATTCCTTCTTAAAT 2563  
DB 2461 TCCCAGATGAAATGTAATTTACTTGGAAAGTAACTTTGGAAGAAATTCCTTCTTAAAT 2520  
QY 2564 CAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2623  
DB 2521 CAAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAACAAAAC 2580

QY 2624 AATTCCTTCATTTTAGCAGTGATGATATGCATAAGTCTGTAAAGCTTGTAACTGGGAAA 2683  
DB 2581 AATTCCTTCATTTTAGCAGTGATGATATGCATAAGTCTGTAAAGCTTGTAACTGGGAAA 2640  
QY 2684 TATTCACCTGATTAATAGCCAGATCTTACTGTATTCCTCCAAAGGCAATATTAGGTAGA 2743  
DB 2641 TATTCACCTGATTAATAGCCAGATCTTACTGTATTCCTCCAAAGGCAATATTAGGTAGA 2700  
QY 2744 TAGATGATTAGTAGTATATTGTTACACACTATTTTGAATTTAGAGAACATACAGAAAGAA 2803  
DB 2701 TAGATGATTAGTAGTATATTGTTACACACTATTTTGAATTTAGAGAACATACAGAAAGAA 2760  
QY 2804 TTTAGGGGCTTAAACATTTAGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATAT 2863  
DB 2761 TTTAGGGGCTTAAACATTTAGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATAT 2820  
QY 2864 TTTTAAATGAATACCAATTTAAATTTTGTATTTTACCTGTTTAAAGAGATTTTGTATCTT 2923  
DB 2821 TTTTAAATGAATACCAATTTAAATTTTGTATTTTACCTGTTTAAAGAGATTTTGTATCTT 2880  
QY 2924 TAAATTTTGTAGTTAAATTTTCTGTGTGATATATATGAGGAATTTACTACTTTATGTC 2983  
DB 2881 TAAATTTTGTAGTTAAATTTTCTGTGTGATATATATGAGGAATTTACTACTTTATGTC 2940  
QY 2984 CTGCTCTTAAACTACATCTGAACTCGACCTCGAGTCTGAGGTATAATAACAAGAGCACTTT 3043  
DB 2941 CTGCTCTTAAACTACATCTGAACTCGACCTCGAGTCTGAGGTATAATAACAAGAGCACTTT 3000  
QY 3044 TTGAGGCAATTTGAAAACCAACCTACACTCTTGGGTCTTAGAGAGATCTGCTGTCTCC 3103  
DB 3001 TTGAGGCAATTTGAAAACCAACCTACACTCTTGGGTCTTAGAGAGATCTGCTGTCTCC 3060  
QY 3104 AAATAAGCTTTTGTATCTGCGAGTGAATTTACTGTAATTAATTTAGTAAATGCTTTCTTTCT 3163  
DB 3061 AAATAAGCTTTTGTATCTGCGAGTGAATTTACTGTAATTAATTTAGTAAATGCTTTCTTTCT 3120  
QY 3164 GGTGATATCTGCTTCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3223  
DB 3121 GGTGATATCTGCTTCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180  
QY 3224 ATCACTGTCTCCCATCTTCCGTGTTAGAGCAAGTGAAGAGTAAAGAGGAGAAAGAA 3283  
DB 3181 ATCACTGTCTCCCATCTTCCGTGTTAGAGCAAGTGAAGAGTAAAGAGGAGAAAGAA 3240  
QY 3284 AGAAGCTCTTACACCACTTTGAGCTCAGACCTCTTAACCCCTGTATTTCCCTTATGATGTC 3343  
DB 3241 AGAAGCTCTTACACCACTTTGAGCTCAGACCTCTTAACCCCTGTATTTCCCTTATGATGTC 3300  
QY 3344 CCCTTTTGGAGACACTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3403  
DB 3301 CCCTTTTGGAGACACTAATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3360  
QY 3404 ACAGTATCTCAGGTGAAATTAACCAACTATAGGCTTTTCTGGGATGATTTCTTA 3463  
DB 3361 ACAGTATCTCAGGTGAAATTAACCAACTATAGGCTTTTCTGGGATGATTTCTTA 3420  
QY 3464 GTCCTTAAAGGTTTGGGACATTTAAACCTTGAGTACATTTGTTGACACAGTTGATTTCC 3523  
DB 3421 GTCCTTAAAGGTTTGGGACATTTAAACCTTGAGTACATTTGTTGACACAGTTGATTTCC 3480  
QY 3524 AAATTTGATGATGGGAGAGAGTCTCTTAAGCTGAGCTTTTCTTGTACTGCAAT 3583  
DB 3481 AAATTTGATGATGGGAGAGAGTCTCTTAAGCTGAGCTTTTCTTGTACTGCAAT 3540  
QY 3584 TATAGAGATTTAGCTTTTAAATTTTGTAGAGTCTAAACACTCTGCTTTCTTACTGCTTA 3643  
DB 3541 TATAGAGATTTAGCTTTTAAATTTTGTAGAGTCTAAACACTCTGCTTTCTTACTGCTTA 3600  
QY 3644 CCTAGTCTGAAACATTTTAAATTAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3703  
DB 3601 CCTAGTCTGAAACATTTTAAATTAAGATTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3660  
QY 3704 AA 3705

Db 3661 AA 3662

RESULT 3

AAAC77143

ID AAC77143 standard; cDNA; 3711 BP.

XX AC AAC77143;

XX AC AAC77143;

DT 08-FEB-2001 (first entry)

XX Human ORFX ORF2698 polynucleotide sequence SEQ ID NO:5395.

DE Human; open reading frame; ORFX; detection; cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant; immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antiviral; antibacterial; antifungal; antirheumatic; antithyroid; antianaemic; gene therapy; cancer; proliferative disorder; hypertensive; neurodegenerative disorder; osteoarthritis; graft vs host disease; cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS; cholesterol ester storage; systemic lupus erythematosus; infection; severe combined immunodeficiency; malaria; autoimmune disorder; asthma; allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound; bone damage; cartilage damage; antiinflammatory disease; coagulation; thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200059473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000MO-US08621.

XX 31-MAR-1999; 99US-0127607.

PR 02-APR-1999; 99US-0127636.

PR 05-APR-1999; 99US-0127728.

PR 30-MAR-2000; 2000US-0540763.

XX (CURA-) CURAGEN CORP.

XX Shinkets RA, Leach M;

XX WPI: 2000-602362/57.

DR P-PSDB; AAB42934.

XX Novel nucleic acids and peptides derived from open reading frame X, useful for treating e.g. cancers, proliferative disorders, neurodegenerative disorders and cardiovascular disease.

PS Claim 5; Page 4578-4580; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397, which represent the human ORFX open reading frames 1 to 3161. The ORFX sequences have activities such as: cytostatic; hepatotropic; vulnary; antipsoriatic; antiparkinsonian; neurotropic; neuroprotective; osteopathic; anticonvulsant; antiarthritic; immunosuppressant; immunostimulant; cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive; dermatological; immunosuppressive; antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic; antithyroid; and antianaemic. The sequences can be used for determining the presence of or predisposition to, or preventing or treating pathological conditions associated with an ORFX-associated disorder. The nucleic acids can be used to express ORFX proteins in gene therapy vectors. The proteins and nucleic acids may be used to treat cancers, proliferative disorders, neurodegenerative disorders, osteoarthritis, graft vs host disease, cardiovascular disease, diabetes mellitus, hypertension, hypothyroidism, cholesterol ester storage, systemic lupus erythematosus, severe combined immunodeficiency (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune disorders, asthma,

CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage, CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance CC coagulation; to inhibit thrombosis; and as a contraceptive.

XX

SQ Sequence 3711 BP; 1215 A; 747 C; 830 G; 919 T; 0 other;

Query Match 92.1%; Score 3412; DB 21; Length 3711;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3702; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 GGCCGACGAGCAGTGGTGTAGCAGCTTGGTGGCAGACAGGTGCGGTAGGTAGAGCGCC 60

DB 6 GGCCGACGAGCAGTGGTGTAGCAGCTTGGTGGCAGACAGGTGCGGTAGGTAGAGCGCC 65

QY 61 GGGACCTGTGACAGGGCTGGTAGCAGCGCAGAGAAAGGGCGCTTTTAGCCAGGTATTTC 120

DB 66 GGGACCTGTGACAGGGCTGGTAGCAGCGCAGAGAAAGGGCGCTTTTAGCCAGGTATTTC 125

QY 121 AGTGTCTGTAGACAAAGATGGAATCATCTCCATTTAATAGACGGCAATGGACCTCAGTATC 180

DB 126 AGTGTCTGTAGACAGGATGGAATCATCTCCATTTAATAGACGGCAATGGACCTCAGTATC 185

QY 181 ATTGAGGTTAACGCCAAAGAACTTCTCTGTCAACAGAAACAAGTCATCGGCTATTGT 240

DB 186 ATTGAGGTTAACGCCAAAGAACTTCTCTGTCAACAGAAACAAGTCATCGGCTATTGT 245

QY 241 GGAATAATTCTCAAGTACCAGAAAGCAGCTGAAGAAACAAACATGGAGAAGAGAG 300

DB 246 GGAATAATTCTCAAGTACCAGAAAGCAGCTGAAGAAACAAACATGGAGAAGAGAG 305

QY 301 TAACACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCTCAGCTGTGTAAAGAGAA 360

DB 306 TAACACCGAAATCTCTCCAGCACTTTAGAAAGGGGACCTCAGCTGTGTAAAGAGAA 365

QY 361 GTGGAGAACCCAGGGCTGGGACGAGTCTCACAGAGCTCTCTACGGAACAGCAGCAC 420

DB 366 GTGGAGAACCCAGGGCTGGGACGAGTCTCACAGAGCTCTCTACGGAACAGCAGCAC 425

QY 421 TGAGATTAGCAGCAGACGACCATCTCTCTGTGAAGTGAAGCAGCAGCTCTCTGG 480

DB 426 TGAGATTAGCAGCAGACGACCATCTCTCTGTGAAGTGAAGCAGCAGCTCTCTGG 485

QY 481 AGCCAAAGTGCACCAAGAAAGCAAAATCCACCCAGATCTAGACTCAGGTCACTCTGA 540

DB 486 AGCCAAAGTGCACCAAGAAAGCAAAATCCACCCAGATCTAGACTCAGGTCACTCTGA 545

QY 541 AGCCCTCGTTTCAAGGTTCGATATCCCCACATCAAGGACGGTGAGGATCTTAAAGACCACTC 600

DB 546 AGCCCTCGTTTCAAGGTTCGATATCCCCACATCAAGGACGGTGAGGATCTTAAAGACCACTC 605

QY 601 AACAGAAAGTAAAAAATCGAAAAATTTGTCTAGGAGAAATCCAGGCATGAAAGTAAAAATC 660

DB 606 AACAGAAAGTAAAAAATCGAAAAATTTGTCTAGGAGAAATCCAGGCATGAAAGTAAAAATC 665

QY 661 AGAAATCAGTGAACACACAGATGCTTCGGGCAAAATAGAGAAATATAATGTTCCTCGTAA 720

DB 666 AGAAATCAGTGAACACACAGATGCTTCGGGCAAAATAGAGAAATATAATGTTCCTCGTAA 725

QY 721 CAGGCTTAAGATGATGTTTGAAGAGGTGAACCACTCAAACTAAGATCTCCGGGCCCA 780

DB 726 CAGGCTTAAGATGATGTTTGAAGAGGTGAACCACTCAAACTAAGATCTCCGGGCCCA 785

QY 781 AAGCCGAAAGTCAAGTGAAGGAGAGATCTCTGAAACAGCTATTCTCTAGATGACCTCGA 840

DB 786 AAGCCGAAAGTCAAGTGAAGGAGAGATCTCTGAAACAGCTATTCTCTAGATGACCTCGA 845

QY 841 AATAGCCAGGTGATGTTTCAATTTTACATTTGACTCGGAGAAAAATGAGATGACG 900

DB 846 AATAGCCAGGTGATGTTTCAATTTTACATTTGACTCGGAGAAAAATGAGATGACG 905

QY 901 AAATCTGGAAGTTCACGCCCTCTCAGAAACCTCTATAAGGATCGAATGCCCAAGTACCA 960

DB 906 AAATCTGGAAGTTCACGCCCTCTCAGAAACCTCTATAAGGATCGAATGCCCAAGTACCA 965

QY 961 GGCAGCTGTGTCCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1020  
DB |||||  
QY 966 GGCAGCTGTGTCCTCAAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGG 1025  
DB |||||  
QY 1021 TGGCGAATCAAAATTCATAAATGGAGCAAAAGAGAAATGTGCCCCCAGTCTCTGAGGT 1080  
DB |||||  
QY 1026 TGGCGAATCAAAATTCATAAATGGAGCAAAAGAGAAATGTGCCCCCAGTCTCTGAGGT 1085  
DB |||||  
QY 1081 CTGCATACCCATCAGGAAGGGGAAAGATTTCTGCAATGAGAAATAGCTGGCAGTCCG 1140  
DB |||||  
QY 1086 CTGCATACCCATCAGGAAGGGGAAAGATTTCTGCAATGAGAAATAGCTGGCAGTCCG 1145  
DB |||||  
QY 1141 TTCCACCCCTGCGGAAGATGATCC...GTGACTCCCGAGTTAAGAGTGAAGTTCAACA 1197  
DB |||||  
QY 1146 TTCCACCCCTGCGGAAGATGATCCCGAGTGAAGTTAAGAGTGAAGTTCAACA 1205  
DB |||||  
QY 1158 GCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCCTCCAGTCTTTCTGAAAGTTC 1257  
DB |||||  
QY 1206 GCCTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCCTCCAGTCTTTCTGAAAGTTC 1265  
DB |||||  
QY 1258 TCCTCCCAAGCAATGAAGAGTTTCAGGCACTTCGCAAGAGAGACCTGCGTGGNATGCA 1317  
DB |||||  
QY 1266 TCCTCCCAAGCAATGAAGAGTTTCAGGCACTTCGCAAGAGAGACCTGCGTGGNATGCA 1325  
DB |||||  
QY 1318 GAAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACAGCAGGTGTTTCACATCAGCTG 1377  
DB |||||  
QY 1326 GAAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACAGCAGGTGTTTCACATCAGCTG 1385  
DB |||||  
QY 1378 CTTCGGTGTCTCTATTTGCAACAAACAACTCAGTCTAGGAACATATGCACTTTTACATGG 1437  
DB |||||  
QY 1386 CTTCGGTGTCTCTATTTGCAACAAACAACTCAGTCTAGGAACATATGCACTTTTACATGG 1445  
DB |||||  
QY 1438 AAGAACTCTATTGAAGCTCAGTCTCACTCACTTTTAAATCTAAGGGCAACTATGATGA 1497  
DB |||||  
QY 1446 AAGAACTCTATTGAAGCTCAGTCTCACTCACTTTTAAATCTAAGGGCAACTATGATGA 1505  
DB |||||  
QY 1498 AGGCTTTGGGCACAGACCAACAGGATCTATGGCAAGCAAAATGAAACCAAGAGAT 1557  
DB |||||  
QY 1506 AGGCTTTGGGCACAGACCAACAGGATCTATGGCAAGCAAAATGAAACCAAGAGAT 1565  
DB |||||  
QY 1558 TTTGGAGAGACCGCCAGCTTGCATAATGCAAGGAGACCCCTCAGAGCCAGGGGTAGA 1617  
DB |||||  
QY 1618 AGATGCCCTATTGCTAAGTGGGTGCTGGCTGCAAGTATGGAAGCCAGCCCTCCTC 1677  
DB |||||  
QY 1626 AGATGCCCTATTGCTAAGTGGGTGCTGGCTGCAAGTATGGAAGCCAGCCCTCCTC 1685  
DB |||||  
QY 1678 TCAGCAGGAGAGCAAGCCAGCTGAACCAAGAGAGCTGAGATCGCTGGCCACC 1737  
DB |||||  
QY 1686 TCAGCAGGAGAGCAAGCCAGCTGAACCAAGAGAGCTGAGATCGCTGGCCACC 1745  
DB |||||  
QY 1738 CCCACTGAACCTTGAAGTTTCAAGAAAGTGCCTTGGAGGAAGGATCAAAATGTCAAAGCC 1797  
DB |||||  
QY 1746 CCCACTGAACCTTGAAGTTTCAAGAAAGTGCCTTGGAGGAAGGATCAAAATGTCAAAGCC 1805  
DB |||||  
QY 1798 CAAATGGCTCTGAAGAGCAAAATCAGAACGCCGAAGTTCTCTGAGGATGTCGATCTAGA 1857  
DB |||||  
QY 1806 CAAATGGCTCTGAAGAGCAAAATCAGAACGCCGAAGTTCTCTGAGGATGTCGATCTAGA 1865  
DB |||||  
QY 1858 TCTGAAGAAGCTAAGACGATCTCTTCACTGAAGGAAGAACGCCCATTCACCTGTAGC 1917  
DB |||||  
QY 1866 TCTGAAGAAGCTAAGACGATCTCTTCACTGAAGGAAGAACGCCCATTCACCTGTAGC 1925  
DB |||||  
QY 1918 AGCTTCATTCAAAGCACCCTGTCTCAAGAGCCCAAAAAGTGTGTCGCCACCTATCAGGAA 1977  
DB |||||  
QY 1926 AGCTTCATTCAAAGCACCCTGTCTCAAGAGCCCAAAAAGTGTGTCGCCACCTATCAGGAA 1985  
DB |||||  
QY 1978 AGGCTGGAGCATGTGAGAGCAGAGTGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAA 2037  
DB |||||  
QY 1986 AGGCTGGAGCATGTGAGAGCAGAGTGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAA 2045  
DB |||||

QY 2038 ACAAGTGGAAATGCCAAGGCTTCTTAAGAAAGATGGAAATGTGGGAAAAACAACCTGGCA 2097  
DB |||||  
QY 2046 ACAAGTGGAAATGCCAAGGCTTCTTAAGAAAGATGGAAATGTGGGAAAAACAACCTGGCA 2105  
DB |||||  
QY 2098 AAACAAAGAAATCTAAAGGAGAGACAGGGAAGAGAAGTAAAGAGGTCATAGTTTGGAGAT 2157  
DB |||||  
QY 2106 AAACAAAGAAATCTAAAGGAGAGACAGGGAAGAGAAGTAAAGAGGTCATAGTTTGGAGAT 2165  
DB |||||  
QY 2158 GGAGAAATGAGAAATCTTTGTAGAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCT 2217  
DB |||||  
QY 2166 GGAGAAATGAGAAATCTTTGTAGAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCT 2225  
DB |||||  
QY 2218 CAACAAACAAATCTCCACAGAACCCCAAGTCTCTGAAATGGTGCAGTTTGTAGACAACAC 2277  
DB |||||  
QY 2226 CAACAAACAAATCTCCACAGAACCCCAAGTCTCTGAAATGGTGCAGTTTGTAGACAACAC 2285  
DB |||||  
QY 2278 CTTTGTGTAAGAAATTCACACTACAGATCAGAAATCCCAAGGATGTGGAACCTCTGGAGGG 2337  
DB |||||  
QY 2286 CTTTGTGTAAGAAATTCACACTACAGATCAGAAATCCCAAGGATGTGGAACCTCTGGAGGG 2345  
DB |||||  
QY 2338 AGAAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAAAATCGGTATATGATGA 2397  
DB |||||  
QY 2346 AGAAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAAAAATCGGTATATGATGA 2405  
DB |||||  
QY 2398 GGATGAGGATGAAGAGTGAACAAATTCGAATGATGTGGCCCTTAAATTCATGTTAGTGT 2457  
DB |||||  
QY 2406 GGATGAGGATGAAGAGTGAACAAATTCGAATGATGTGGCCCTTAAATTCATGTTAGTGT 2465  
DB |||||  
QY 2458 AGCGAGCCACTGCCCTTTGTGCAAAATGTGATGCAATTAAGCAGGTATCCAGCATGAAT 2517  
DB |||||  
QY 2466 AGCGAGCCACTGCCCTTTGTCAAAATGTGATGCAATAAGCAGGTATCCAGCATGAAT 2525  
DB |||||  
QY 2518 GTAATTTACTTGGAAAGTAACTTTGGAAGAAATTCCTTTTAAATCAAAAACAACAA 2577  
DB |||||  
QY 2526 GTAATTTACTTGGAAAGTAACTTTGGAAGAAATTCCTTTTAAATCAAAAACAACAA 2585  
DB |||||  
QY 2578 AAAAAACAAAAACACATTTCTAAATCTAGAGATAAAGTAACTTAAATCTTCAATTTA 2637  
DB |||||  
QY 2586 AAAAAACAAAAACACATTTCTAAATCTAGAGATAAAGTAACTTAAATCTTCAATTTA 2645  
DB |||||  
QY 2638 GCAGTGTATGATGATGAAGTCTGTAAAGGTTGTAACTGGGGAATAATTCACCTGATA 2697  
DB |||||  
QY 2646 GCAGTGTATGATGATGAAGTCTGTAAAGGTTGTAACTGGGGAATAATTCACCTGATA 2705  
DB |||||  
QY 2698 ATAGCCAGATTTCTACTGTATTTCCAAAAGGCAATATTAAGGTAGATAGATGATTAGTAG 2757  
DB |||||  
QY 2706 ATAGCCAGATTTCTACTGTATTTCCAAAAGGCAATATTAAGGTAGATAGATGATTAGTAG 2765  
DB |||||  
QY 2758 TATATTGTTACACACTATTTTGGAAATTAGAGAACATACAGAGGAAATTTAGGGCTTAAA 2817  
DB |||||  
QY 2766 TATATTGTTACACACTATTTTGGAAATTAGAGAACATACAGAGGAAATTTAGGGCTTAAA 2825  
DB |||||  
QY 2818 CATTACGACTGAATGCACTTTTAGTATAAAGGGCACAGTTTGTATATTTTAAATGAATAC 2877  
DB |||||  
QY 2826 CATTACGACTGAATGCACTTTTAGTATAAAGGGCACAGTTTGTATATTTTAAATGAATAC 2885  
DB |||||  
QY 2878 CAATTTAAATTTTGTAGTATTTTACCTGTTAAGAGATTTTGTAGTCTTTTAAATTTT 2937  
DB |||||  
QY 2886 CAATTTAAATTTTGTAGTATTTTACCTGTTAAGAGATTTTGTAGTCTTTTAAATTTT 2945  
DB |||||  
QY 2938 TAATTTTCTTGTGTATATATATAGGAAATTTTACTTTTATGTCTGTCTCTTAACT 2997  
DB |||||  
QY 2946 TAATTTTCTTGTGTATATATATAGGAAATTTTACTTTTATGTCTGTCTCTTAACT 3005  
DB |||||  
QY 2998 ACATCTGAACTCGAGCTCTGAGGTATATAACAGAGCACCTTTTGGGCAATGAA 3057  
DB |||||  
QY 3006 ACATCTGAACTCGAGCTCTGAGGTATATAACAGAGCACCTTTTGGGCAATGAA 3065  
DB |||||  
QY 3058 AAACCAACCTTACACTCTTGGGTCTTAGAGAGATCTGCTCTCCCAATAAGCTTTTGT 3117  
DB |||||  
QY 3066 AAACCAACCTTACACTCTTGGGTCTTAGAGAGATCTGCTCTCCCAATAAGCTTTTGT 3125  
DB |||||  
QY 3118 ATCTGCCAGTAATTTACTGTACTCCAAATGATGCTTTCTTTCTGTGTATCTGTGC 3177  
DB |||||

Db 3126 ATCTGCCAGTGAATTTACTGTACTCCAAATGATGCTTCTTTCTGGTGATCTGTGC 3185  
QY TTCTCATAATTACTGAAGCTCAATATTTTAGTATACTTCGGGATCACTGTCCCCCA 3237  
Db TTCTCATAATTACTGAAGCTCAATATTTTAGTATACTTCGGGATCACTGTCCCCCA 3245  
QY TCTTCCGTGTTAGAGCAAAAGTGAAGAGTTTAAAGGAGGAGAAAGAACTGCTTACA 3297  
Db TCTTCCGTGTTAGAGCAAAAGTGAAGAGTTTAAAGGAGGAGAAAGAACTGCTTACA 3305  
QY CACATTGAGCTCAGACTCTAAACCCCTGATTTCCCTTATGATGCCCTTTTTCAGACA 3357  
Db CCACATTGAGCTCAGACTCTAAACCCCTGATTTCCCTTATGATGCCCTTTTTCAGACA 3365  
QY CTAAATTTTAAATACTACTAGCTCTGAATATATGATTTTATACAGATTTCTCAGG 3417  
Db CTAATTTTAAATACTACTAGCTCTGAATATATGATTTTATACAGATTTCTCAGG 3425  
QY GTGAAATTAAACCACTATAGGCTTTTCTTTGGGATGATTTTCTAGTCTTAAGSTTTGG 3477  
Db GTGAAATTAAACCACTATAGGCTTTTCTTTGGGATGATTTTCTAGTCTTAAGSTTTGG 3485  
QY GGACATTATAAATTGAGTACATTTGTTGTACACAGTTGATATTCATTCATTTGATGATG 3537  
Db GGACATTATAAATTGAGTACATTTGTTGTACACAGTTGATATTCATTCATTTGATGATG 3545  
QY GGAGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGGGATGATTTTCTAGTCTTAAGSTTTAGC 3597  
Db GGAGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGGGATGATTTTCTAGTCTTAAGSTTTAGC 3605  
QY TTTAATATTTTTTAGAGATGTAACACATTTCTGCTTTCTTTAGTCTTAAGSTTTAGC 3657  
Db TTTAATATTTTTTAGAGATGTAACACATTTCTGCTTTCTTTAGTCTTAAGSTTTAGC 3665  
QY TTTTATTCAATAAGATTTTAAATTAATTTGAAAAATTTGAAAAATTTGAAAAATTTGAAAA 3703  
Db TTTTATTCAATAAGATTTTAAATTAATTTGAAAAATTTGAAAAATTTGAAAAATTTGAAAA 3711

RESULT 4

AAZ33566  
ID AAZ33566 standard; cDNA; 2783 BP.  
XX  
AC AAZ33566;  
XX  
DT 08-DEC-1999 (first entry)  
XX  
DE Human breast tumour-associated EST 26.  
XX  
KW Expressed sequence tag; EST; human; breast; cancer; cytostatic;  
KW medicaments; gene therapy; treatment; fat metabolism; ss.  
XX  
OS Homo sapiens.  
XX  
PN DE19813835-A1.  
XX  
PD 23-SEP-1999.  
XX  
PF 20-MAR-1998; 98DE-1013835.  
XX  
PR 20-MAR-1998; 98DE-1013835.  
XX  
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.  
XX  
PI Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;  
XX  
DR WPI; 1999-528979/45.  
DR P-PSDB; AAY48487.  
XX  
PT Human nucleic acid sequences and protein products from normal breast  
XX tissue, useful for breast cancer therapy

PS Claim 1a; 113-114; 206pp; German.  
XX This invention describes novel human nucleic acid sequences from normal breast tissue which have cytostatic activity. The nucleic acid sequences can be used to produce and isolate full-length gene sequences. They can be used to express proteins, which can be used as tools to find an activity against breast cancer. The sequences can be used in sense or antisense form. They are especially useful for medicaments for gene therapy to treat breast cancer and for treating illnesses associated with fat metabolism. AAZ33541-33610 represent expressed sequence tags described in the method of the invention.  
SQ Sequence 2783 BP; 901 A; 527 C; 608 G; 747 T; 0 other;  
Query Match 70.1%; Score 2596; DB 20; Length 2783;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2696; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1001 AATGAGCTGAAGCCAGTGTGGCGAAATCAAAATTCATAAAATGGAGCAAAAGGAGAAT 1060  
Db 40 AATGAGCTGAAGCCAGTGTGGCGAAATCAAAATTCATAAAATGGAGCAAAAGGAGAAT 99  
QY 1061 GTGCCCCCAGGTCTCTGAGTCTGCATCAATCAGGAAAGGAAAGATTTCTGCAAAAT 1120  
Db 100 GTGCCCCCAGGTCTCTGAGTCTGCATCAATCAGGAAAGGAAAGATTTCTGCAAAAT 159  
QY 1121 GAGAATAGCTGCGAGTCTCCACCCCTCCGCAAGATGACTCCCGTACTCCAGGTT 1180  
Db 160 GAGAATAGCTGCGAGTCTCCACCCCTCCGCAAGATGACTCCCGTACTCCAGGTT 219  
QY 1181 AAGAGTGAAGTTCACAGCGCTGCCATCCAGCCACTAAGTCCAGATTCAGAGGCTCC 1240  
Db 220 AAGAGTGAAGTTCACAGCGCTGCCATCCAGCCACTAAGTCCAGATTCAGAGGCTCC 279  
QY 1241 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAGTTTCAGGCACCTGCAAGAGAG 1300  
Db 280 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAGTTTCAGGCACCTGCAAGAGAG 339  
QY 1301 ACTGCTGGAATGTCAGAAAGCAGTCTATCAATGGAGCGTCTCTTGCCCAACAGCAG 1360  
Db 340 ACTGCTGGAATGTCAGAAAGCAGTCTATCAATGGAGCGTCTCTTGCCCAACAGCAG 399  
QY 1361 GTGTTTTCATCAGCTGCTTCGTTGCTCTTATGCAACAAACAACTAGTCTAGGAACA 1420  
Db 400 GTGTTTTCATCAGCTGCTTCGTTGCTCTTATGCAACAAACAACTAGTCTAGGAACA 459  
QY 1421 TATGATCTTTTACATGGAAGATCTATTGTAAGCTCACTTCAATCAACTCTTTAAATCT 1480  
Db 460 TATGATCTTTTACATGGAAGATCTATTGTAAGCTCACTTCAATCAACTCTTTAAATCT 519  
QY 1481 AAGGCAACTATGATGAAGGCTTTGGCACAGACACACAAGGATCTATGGCAAGCAAA 1540  
Db 520 AAGGCAACTATGATGAAGGCTTTGGCACAGACACACAAGGATCTATGGCAAGCAAA 579  
QY 1541 AATGAAAACGAGAGATTTTGGAGAGACCGCCAGCTTGCAAAATGCAAGGAGACCCCT 1600  
Db 580 AATGAAAACGAGAGATTTTGGAGAGACCGCCAGCTTGCAAAATGCAAGGAGACCCCT 639  
QY 1601 CACAGCCAGGGGTAGAAGATGCCCCCTATTGCTAAGGTGGGTCTCTGGCTGCAAGTATG 1660  
Db 640 CACAGCCAGGGGTAGAAGATGCCCCCTATTGCTAAGGTGGGTCTCTGGCTGCAAGTATG 699  
QY 1661 GAAGCCAAAGGCTCTCTCAGCAGGAGAAAGAACAGCCAGCTGAAACCAAGAAAGCTG 1720  
Db 700 GAAGCCAAAGGCTCTCTCAGCAGGAGAAAGAACAGCCAGCTGAAACCAAGAAAGCTG 759  
QY 1721 AGGATCCCTGGCCACCCCTCACTGAATCTTGAAGTTCAGGAAGTGCCTTGGAGAAAGGG 1780  
Db 760 AGGATCCCTGGCCACCCCTCACTGAATCTTGAAGTTCAGGAAGTGCCTTGGAGAAAGGG 819  
QY 1781 ATCAAAATGTCAAAGCCCAATGSCCTCTGAAAGCAGAAATCAGCAAGCCGAGTTTCT 1840  
Db 820 ATCAAAATGTCAAAGCCCAATGSCCTCTGAAAGCAGAAATCAGCAAGCCGAGTTTCT 879





[illegible]







QY 2182 TGGTCAGACTCCGATGAAGATGATAACAGCTTCTCTCAAAACAAATCTCCACAGAACC 2241  
Db 2218 TGGTCAGACTCCGATGAAGATGATAACAGCTTCTCTCAAAACAAATCTCCACAGAACC 2277  
QY 2242 CAAGTCTCTGAATTCGGTCGAGTTTGTGTAGACACACCTTTCCTGAAGAAATTCATCTCA 2301  
Db 2278 CAAGTCTCTGAATTCGGTCGAGTTTGTGTAGACACACCTTTCCTGAAGAAATTCATCTCA 2337  
QY 2302 GAATCAGAAATCCAGGATGTGGAATCTCTGGAGGAGAGAGTGGTCAAGAGCTCTCTGT 2361  
Db 2338 GAATCAGAAATCCAGGATGTGGAATCTCTGGAGGAGAGAGTGGTCAAGAGCTCTCTGT 2397  
QY 2362 GGAAGACAGATAAAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGCACAAAT 2421  
Db 2398 GGAAGACAGATAAAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGCACAAAT 2457  
QY 2422 TGCATGATGCTGGCCCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAA 2481  
Db 2458 TGCATGATGCTGGCCCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAA 2517  
QY 2482 ATGTGATGCACATAGCAGGATATCCAGCATGAATGTAAATTTACTTTGGAAGTAACTTTG 2541  
Db 2518 ATGTGATGCACATAGCAGGATATCCAGCATGAATGTAAATTTACTTTGGAAGTAACTTTG 2577  
QY 2542 GAAAGAAATTCCTTTTAAATTCAAAAACAAAAACAAAAACACA 2586  
Db 2578 GAAAGAAATTCCTTTTAAATTCAAAAACAAAAACAAAAACACA 2622

RESULT 7  
AAF55697  
ID AAF55697 standard; DNA, 3650 BP.  
XX  
AC AAF55697;  
XX  
DT 11-JUN-2001 (first entry)  
XX  
DE Human EPLIN (epithelial protein lost in neoplasm)-beta isoform DNA.  
KW Human; EPLIN; epithelial protein lost in neoplasm; EPLIN-alpha;  
KW EPLIN-beta; tumour suppressor; tumour; cell proliferative disorder;  
KW gene therapy; cancer; ss.  
XX  
OS Homo sapiens.

Key Location/Qualifiers  
FT 102..2384  
FT CDS  
FT /\*tag= a  
FT /transl\_except= "{pos: 933..935, aa: Val}"  
FT /transl\_except= "{pos: 1131..1136, aa: Arg}"  
FT /transl\_except= "{pos: 1587..1589, aa: Asp}"  
FT /transl\_except= "{pos: 1635..1637, aa: Lys}"  
FT /transl\_except= "{pos: 2068..2070, aa: Ser}"  
FT /transl\_except= "{pos: 2076..2078, aa: Ser}"  
FT /transl\_except= "{pos: 2139..2141, aa: Ile}"  
FT /transl\_except= "{pos: 2196..2198, aa: Ser}"  
FT /transl\_except= "{pos: 2214..2216, aa: Ser}"  
FT /product= "EPLIN (epithelial protein lost in neoplasm)"

PN WO200119019-A1.  
XX  
PD 15-MAR-2001.  
XX  
PD 08-SEP-2000; 2000WO-US24689.  
XX  
PR 08-SEP-1999; 99US-0153024.  
XX (REGC ) UNIV CALIFORNIA.  
XX  
PI Chang DD, Maul RS;  
XX WPI; 2001-244555/25.  
DR

DR P-PSDB; AAB67701.  
XX  
PT New tumor suppressor protein EPLIN, useful as a marker for diagnostic,  
PT prognostic and therapeutic applications over the course of cell  
XX proliferative disorders associated with EPLIN .  
PS Claim 4; Page 43-44; 59pp; English.  
XX  
CC The present sequence encodes a human EPLIN (epithelial protein lost in  
CC neoplasm)-beta isoform. The specification also describes EPLIN-alpha.  
CC EPLIN is a tumour suppressor protein, whose expression is altered in  
CC multiple common human tumour types. EPLIN nucleic acids and proteins are  
CC used in screening assays to detect molecules that specifically bind to  
CC EPLIN nucleic acids, proteins or derivatives and thus have potential use  
CC as agonist or antagonist of EPLIN, in particular molecules that affect  
CC cell proliferation. Thus the assays are useful for screening molecules  
CC with potential utility as anticancer drugs or lead compounds for drug  
CC development. EPLIN nucleic acids, proteins are useful for detecting a  
CC cell proliferative disorder in a subject. EPLIN polynucleotides are  
CC useful in gene therapy techniques. EPLIN is useful as a marker that  
CC can be diagnostically, prognostically and therapeutically used over  
CC the course of a cell proliferative disorder associated with EPLIN.  
XX

SQ Sequence 3650 BP; 1197 A; 733 C; 809 G; 911 T; 0 other;  
Query Match 57.5%; Score 2130; DB 22; Length 3650;  
Best Local Similarity 99.4%; Pred No. 0;  
Matches 3580; Conservative 0; Mismatches 11; Indels 11; Gaps 4;

QY 92 AGGAAAGCGGCTTTTAGCCAGGTATTTTCAGTGTCTGTAGACAAGATGGAATCATCTCCA 151  
Db 57 AGGAAAGCGGCTTTTAGCCAGGTATTTTCAGTGTCTGTAGACAAGATGGAATCATCTCCA 116  
QY 152 TTTAATAGACGCAATGACCTCCTATCATTTAGGGTAAACAGCCAAAGAACTTTCTCTT 211  
Db 117 TTTAATAGACGCAATGACCTCCTATCATTTAGGGTAAACAGCCAAAGAACTTTCTCTT 176  
QY 212 GTCAACAAGAACAGTCACTCGCTATTTGTGGAATATTTCTCAAGTACAGAAACAGCT 271  
LJ 177 GTCAACAAGAACAGTCACTCGCTATTTGTGGAATATTTCTCAAGTACAGAAACAGCT 236  
QY 272 GAAGAAACAAACATGGGAG 331  
Db 237 GAAGAAACAAACATGGGAG 296  
QY 332 AAGGGACCTTGACTGTGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 391  
Db 297 AAGGGACCTTGACTGTGTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 356  
QY 392 CACACAGACTCTCTACGGAACAGCAGCACTGAGATTAGGCACAGAGCAGACCATCTCTCT 451  
Db 357 CACACAGACTCTCTACGGAACAGCAGCACTGAGATTAGGCACAGAGCAGACCATCTCTCT 416  
QY 452 GCTGAAGTGACAAGCCACGCTGCTTCTGGAGCCAAAGCTACCAAGAGAACAAATCCAC 511  
Db 417 GCTGAAGTGACAAGCCACGCTGCTTCTGGAGCCAAAGCTACCAAGAGAACAAATCCAC 476  
QY 512 CCCAGATCTAGACTCAGGTCACTCTCTGGAAGCCCTCTGTTCCAGGGTGCATATCCCCCATC 571  
Db 477 CCCAGATCTAGACTCAGGTCACTCTCTGGAAGCCCTCTGTTCCAGGGTGCATATCCCCCATC 536  
QY 572 AAGGACGCTGAGGATCTTAAAGACCACTCAACAGAAAGTAAAAAATGGAAAAATTTGCTA 631  
Db 537 AAGGACGCTGAGGATCTTAAAGACCACTCAACAGAAAGTAAAAAATGGAAAAATTTGCTA 596  
QY 632 GGAGAAATCCAGGCATGAAGTAGAAAAATCAGAAATCAGTGAAGAACACAGATGCTCGGGC 691  
Db 597 GGAGAAATCCAGGCATGAAGTAGAAAAATCAGAAATCAGTGAAGAACACAGATGCTCGGGC 656  
QY 692 AAAATAGAGAAATATAATGTTCCGCTGGAACAGGCTTAAGATGATGTTTGAGAAAGTGAA 751  
Db 657 AAAATAGAGAAATATAATGTTCCGCTGGAACAGGCTTAAGATGATGTTTGAGAAAGTGAA 716

QY 752 CCAACTCAAAGTAAAGTTCTCCGGGCCCAAGCCGAAAGTGCAGTGGAAAGCAATCTCT 811  
Db 717 CCAACTCAAAGTAAAGTTCTCCGGGCCCAAGCCGAAAGTGCAGTGGAAAGCAATCTCT 776  
QY 812 GAAAGACGCTATTCTCTAGATGACCTCGGAATAGCCCGCAGTGAAGTGTCTCTTACA 871  
Db 777 GAAAGACGCTATTCTCTAGATGACCTCGGAATAGCCCGCAGTGAAGTGTCTCTTACA 836  
QY 872 TTTGACTCGGAGAAAATGAGAGTAGACGAAATCTGGAAATCTCCAGCCCTCTCAGAAACC 931  
Db 837 TTTGACTCGGAGAAAATGAGAGTAGACGAAATCTGGAAATCTCCAGCCCTCTCAGAAACC 896  
QY 932 TCTATAAGAGATCGAAATGGCCAAAGTACCAGGCAAGTGTGTCCAAACAAAGCAAGCTCAACC 991  
Db 897 TCTATAAGAGATCGAAATGGCCAAAGTACCAGGCAAGTGTGTCCAAACAAAGCAAGCTCAACC 956  
QY 992 AACTATACAAATGAGCTGAAAGCCAGTGGTGGGAAATCAAAATTCATAAAATGGAGCAA 1051  
Db 957 AACTATACAAATGAGCTGAAAGCCAGTGGTGGGAAATCAAAATTCATAAAATGGAGCAA 1016  
QY 1052 AAGGAGAAATGTGCCCCCAGGCTCTGAGTCTGATCAGCCATCAGGAAGGGGAAAGATT 1111  
Db 1017 AAGGAGAAATGTGCCCCCAGGCTCTGAGTCTGATCAGCCATCAGGAAGGGGAAAGATT 1076  
QY 1112 TCTGCAAAATGAGAAATAGCCTGGCAGTCCGTTCCAGCCCTGCCGAGATGACTCCC-CT 1168  
Db 1077 TCTGCAAAATGAGAAATAGCCTGGCAGTCCGTTCCAGCCCTGCCGAGATGACTCCCAGGT 1136  
QY 1169 GACTCCGAGTTAAGAGTGAGTTCAACAGCTGTCCATCCCAAGCCACTAAGTCCAGAT 1228  
Db 1137 GACTCCGAGTTAAGAGTGAGTTCAACAGCTGTCCATCCCAAGCCACTAAGTCCAGAT 1196  
QY 1229 TCCAGAGCCTCCAGTCTTCTGAAAGTCTCTCCCTCCCAAGCAATGAGAGTTTCAGGCA 1288  
Db 1197 TCCAGAGCCTCCAGTCTTCTGAAAGTCTCTCCCTCCCAAGCAATGAGAGTTTCAGGCA 1256  
QY 1289 CCTGCAAGAGACCTGCGTGGAAATGTCAGAGACAGTCTATCCCAATGGAGGCTCTCTTG 1348  
Db 1257 CCTGCAAGAGACCTGCGTGGAAATGTCAGAGACAGTCTATCCCAATGGAGGCTCTCTTG 1316  
QY 1349 GCGAACACAGAGTTTACATCAGTCTGCTTCCGTTGCTCTTATGCAACCAAACTC 1408  
Db 1317 GCGAACACAGAGTTTACATCAGTCTGCTTCCGTTGCTCTTATGCAACCAAACTC 1376  
QY 1409 AGCTAGGAACATATGCACTCTTACATGGAAGATCTATTGTAAGCCTCACTTCAATCAA 1468  
Db 1377 AGCTAGGAACATATGCACTCTTACATGGAAGATCTATTGTAAGCCTCACTTCAATCAA 1436  
QY 1469 CTCTTTAAATCTAAGGGCAACTATGATGAAGCTTTGGGACAGACACAAAGATCTA 1528  
Db 1437 CTCTTTAAATCTAAGGGCAACTATGATGAAGCTTTGGGACAGACACAAAGATCTA 1496  
QY 1529 TGGGCAAGCAAAAATGAAACGAAGAGATTTTGGAGAGACCGCCAGCTTGCAAAATGCA 1588  
Db 1497 TGGGCAAGCAAAAATGAAACGAAGAGATTTTGGAGAGACCGCCAGCTTGCAAAATGCA 1556  
QY 1589 AGGAGAGCCCTCACAGCCCGAGGGTAGAAGATGCCCTTATGTAAGTGGTGTCTTG 1648  
Db 1557 AGGAGAGCCCTCACAGCCCGAGGGTAGAAGATGCCCTTATGTAAGTGGTGTCTTG 1616  
QY 1649 GCTGCAAGTATGGAAGCCAAAGCCCTCTCTCAGCAGGAGAGGAGACCAAGCCAGCTGAA 1708  
Db 1617 GCTGCAAGTATGGAAGCCAGGCTCTCTCTCAGCAGGAGAGGAGACCAAGCCAGCTGAA 1676  
QY 1709 ACCAAGAGCTGAGGATGCGCTGCGCCACCCCACTGAACTTGGAACTTCAGGAAGTGCC 1768  
Db 1677 ACCAAGAGCTGAGGATGCGCTGCGCCACCCCACTGAACTTGGAACTTCAGGAAGTGCC 1736  
QY 1769 TTGAGCAAGGGATCAAAATGTCAAGCCCAATGGCCCTCTGAAAGCAAGATCAGCAAG 1828  
Db 1737 TTGAGCAAGGGATCAAAATGTCAAGCCCAATGGCCCTCTGAAAGCAAGATCAGCAAG 1796  
QY 1829 CCCGAAGTTCCTGAGGATGTCGATCTAGATCTGAAAGCTAAGACGATCTTCTTCACTG 1888

Db 1797 CCCGAAGTTCCTGAGGATGTCGATCTAGATCTGAAGAGCTTAAGACGATCTTCTTCACTG 1856  
QY 1889 AAGGAAAGAGCGGCCATTCTACTGTACGAGTTCATTTCAAAGCACCTCTCTCAAGAGC 1948  
Db 1857 AAGGAAAGAGCGGCCATTCTACTGTAGCAGTTCATTTCAAAGCACCTCTCTCAAGAGC 1916  
QY 1949 CCAAAAACCTGTGTCCCCACCTATCAGGAAAGGCTCGAGCATGTCTAGACGACGAATGAAGAA 2008  
Db 1917 CCAAAAACCTGTGTCCCCACCTATCAGGAAAGGCTCGAGCATGTCTAGACGACGAATGAAGAA 1976  
QY 2009 TCTGTGGTGGAAAGTTGCAAGAAAGGAAACAAGTGGAAAATGCCAAGCTTCTAAGAG 2068  
Db 1977 TCTGTGGTGGAAAGTTGCAAGAAAGGAAACAAGTGGAAAATGCCAAGCTTCTAAGAG 2036  
QY 2069 AATGGGAAATGTGGGAAAACAACCTGGCAAAACAAGAAATTTAAAGGAGACACAGGGAAG 2128  
Db 2037 AATGGGAAATGTGGGAAAACAACCTGGCAAAACAAGAAATTTAAAGGAGACACAGGGAAG 2096  
QY 2129 AGAAGTAAGGAGGTTCATAGTTTGGAGATGGAGATGAGAAATCTGTAGAAAATGGTGCA 2188  
Db 2097 AGAAGTAAGGAGGTTCATAGTTTGGAGATGGAGATGAGAAATTTGTAGAAAATGGTGCA 2156  
QY 2189 GACTCCGATGAAGATGATAACAGCTTCTCAAAACAACATCTCCACAAGAACCCCAAGTCT 2248  
Db 2157 GACTCCGATGAAGATGATAACAGCTTCTCAAAACAACATTTCCACAAGAACCCCAAGTCT 2216  
QY 2249 CTGAATTTGTCGAGTTTGTAGACAAACCTTTGCTGAAGAAATTCACCTCTCAGAAATCAG 2308  
Db 2217 TTGAATTTGTCGAGTTTGTAGACAAACCTTTGCTGAAGAAATTCACCTCTCAGAAATCAG 2276  
QY 2309 AAATCCCAGGATGTGAACCTCTGGGAGGAGAGTGTCAAAGAGCTCTCTGTGGAAAGAA 2368  
Db 2277 AAATCCCAGGATGTGAACCTTTGGGAGGAGAGTGTCAAAGAGCTCTCTGTGGAAAGAA 2336  
QY 2369 CAGATAAGAGAAATCGGTATTATGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 2428  
Db 2337 CAGATAAGAGAAATCGGTATTATGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATG 2396  
QY 2429 ATGCTGGGCTTTAAATTCATGTTAGTGTAGCGAGGACCTGCCCCCTTTGTCAAATGTGAT 2488  
Db 2397 ATGCTGGGCTTTAAATTCATGTTAGTGTAGCGAGGACCTGCCCCCTTTGTCAAATGTGAT 2456  
QY 2489 GCACATTAAGCAGGTATCCCGAGCATGAAATGTAATTTACTTGGAAAGTAACTTTGGAAGAA 2548  
Db 2457 GCACATTAAGCAGGTATCCCGAGCATGAAATGTAATTTACTTGGAAAGTAACTTTGGAAGAA 2516  
QY 2549 ATTCCCTCTTAAATCAAAAACAACCAAAAACAACCAAAAACAACCAAAAACAACCAAAAACA 2608  
Db 2517 ATTCCCTCTTAAATCAAAAACAACCAAAAACAACCAAAAACAACCAAAAACAACCAAAAACA 2576  
QY 2609 AGATAACTTTACTTAAATTTCTTCATTTTAGCAGTGTATGATATGATGATGATGATGATGATGATG 2668  
Db 2577 AGATAACTTTACTTAAATTTCTTCATTTAGCAGTGTATGATGATGATGATGATGATGATGATGATG 2631  
QY 2669 TTGTAACCTGGGAAATATTTCCACCTGATAATAGCCCGAGATTTACTGTATTTCCCAAGG 2728  
Db 2632 TTGTAACCTGGGAAATATTTCCACCTGATAATAGCCCGAGATTTACTGTATTTCCCAAGG 2691  
QY 2729 CAATATTAAGGTAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2788  
Db 2692 CAATATTAAGGTAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2751  
QY 2789 AACATACAGAGAAATTTAGGGCTTTAAACATTCAGCTGAACTGCACTTTAGTATAAAGG 2848  
Db 2752 AACATACAGAGAAATTTAGGGCTTTAAACATTCAGCTGAACTGCACTTTAGTATAAAGG 2810  
QY 2849 GCACAGTTTGTATATTTTAAATGAATACCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2908  
Db 2811 GCACAGTTTGTATATTTTAAATGAATACCAATTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2870  
QY 2909 AGATTATTTAGTCTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 2968

Db 2871 AGATTATTAGTCTTTAAATTTTGTAGTTAAATTTCTTGCTGTGATATATAGGGAAT 2930  
Qy 2959 TTACTACTTTATGCTGCTCTCTAAACTACATCTGAACTCGACGCTCGCTGAGTATAAT 3028  
Db 2931 TTACTACTTTATGCTGCTCTCTAAACTACATCTGAACTCGACGCTCGCTGAGTATA-- 2988  
Qy 3029 ACAACAGAGCACTTTTGAGGCAATTTGAAACCAACTACACTCTTCGGTGTCTTACAGA 3088  
Db 2989 ACAACAGAGCACTTTTGAGGCAATTTGAAACCAACTACACTCTTCGGTGTCTTACAGA 3048  
Qy 3089 CATCTGCTGCTCTCCAAATAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAATG 3148  
Db 3049 CATCTGCTGCTCTCCAAATAGCTTTTGTATCTGCCAGTGAATTTACTGTACTCCAAATG 3108  
Qy 3149 ATTGCTTTCTTTCTGGTGATATCTGCTCTCTCATAAATTTACTGAAAGCTGCAATTTTT 3208  
Db 3109 ATTGCTTTCTTTCTGGTGATATCTGCTCTCTCATAAATTTACTGAAAGCTGCAATTTTT 3168  
Qy 3209 AGTAATACCTTCGGGATCACTGCTCCCATCTTCCTGGTTAGAGCAAGTGAAGAGTTTA 3268  
Db 3169 AGTAATACCTTCGGGATCACTGCTCCCATCTTCCTGGTTAGAGCAAGTGAAGAGTTTA 3228  
Qy 3269 AAGGAGGAAGAAAGAACTGCTTACACCACTTGTAGCTCAGACCTTAAACCCCTGTAT 3328  
Db 3229 AAGGAGGAAGAAAGAACTGCTTACACCACTTGTAGCTCAGACCTTAAACCCCTGTAT 3288  
Qy 3329 TTCCCTTATGATGCTCCCTTTTGTAGACACTAAATTTTAAATCTACTAGCTCTCAAAAT 3388  
Db 3289 TTCCCTTATGATGCTCCCTTTTGTAGACACTAAATTTTAAATCTACTAGCTCTCAAAAT 3348  
Qy 3389 ATATTGATTTTATCAGATTTCTCAGGCTGAAATTTAAACCAACTATAGGCTTTTCT 3448  
Db 3349 ATATTGATTTTATCAGATTTCTCAGGCTGAAATTTAAACCAACTATAGGCTTTTCT 3408  
Qy 3449 TGGGATGATTTTCTAGCTTTAAGGTTTGGGACATTTAACTTGTAGTACATTTGTGTA 3508  
Db 3409 TGGGATGATTTTCTAGCTTTAAGGTTTGGGACATTTAACTTGTAGTACATTTGTGTA 3468  
Qy 3509 CACAGTTGATTTTCCAAATTTGATGATGGGAGAGTGCTCTTAACTGTAGCTTT 3568  
Db 3469 CACAGTTGATTTTCCAAATTTGATGATGGGAGAGTGCTCTTAACTGTAGCTTT 3528  
Qy 3569 TCTTTGATGCTTTATAGAGATTTAGCTTTTAAATTTTATAGATCTAAACATTTCT 3628  
Db 3529 TCTTTGATGCTTTATAGAGATTTAGCTTTTAAATTTTATAGATCTAAACATTTCT 3588  
Qy 3629 GCTTTCTTGTCTTACCTAGCTGTAACATTTTATTCATTAAGATTTTAAATTTAAAT 3688  
Db 3589 GCTTTCTTGTCTTACCTAGCTGTAACATTTTATTCATTAAGATTTTAAATTTAAAT 3648  
Qy 3689 TG 3690  
Db 3649 TG 3650

RESULT 8

AAC59489

ID AAC59489 standard; cDNA; 2267 BP.

XX AAC59489;

AC AAC59489;

DT 26-JAN-2001 (first entry)

XX Human secreted protein gene 41 SEQ ID NO:51.

DE Human;

XX Human; secreted protein; diagnosis; antiarthritic; immunosuppressive;

KW antiarthritic; antiproliferative; cytosolic; cardiast; vasotropic;

KW cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide;

KW fungicide; ophthalmological; vulnary; gene therapy; autoimmune disease;

KW hyperproliferative disorder; neoplasia; cancer; cardiovascular disorder;

KW cerebrovascular disorder; angiogenesis; nervous system disorder;

KW infection; ocular disorder; wound healing; skin aging; food additive;

KW preservative; ss.

XX Homo sapiens.  
OS  
XX  
PN W0200056755-A1.  
XX  
XX 28-SEP-2000.  
XX  
XX 16-MAR-2000; 2000WO-US06930.  
XX  
PR 19-MAR-1999; 99US-0125361.  
PR 10-DEC-1999; 99US-0169910.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Ruben SM, Komatsoulis G;  
XX  
XX WPI; 2000-587661/55.  
DR P-PSDB; AAB34132.  
XX  
XX  
PT New isolated nucleic acid molecules encoding 49 human secreted proteins  
PT used for preventing, treating or ameliorating medical conditions, for  
PT diagnosing pathological conditions or as food additives or  
PT preservatives  
XX  
PS Claim 1; Page 357-358; 419pp; English.  
XX  
XX The polynucleotide sequences given in AAC59449 to AAC59497 encode the  
CC human secreted proteins given in AAB34092 to AAB34140. AAB34141 to  
CC AAB34216 represent human secreted polypeptide sequences and proteins  
CC homologous to them, which are given in the exemplification of the present  
CC invention. Human secreted proteins have activities based on the tissue  
CC and cells the genes are expressed in. Examples of activities include:  
CC antiarthritic; immunosuppressive; antirheumatic; antiproliferative;  
CC cytosolic; cardiast; vasotropic; cerebroprotective; neurotropic;  
CC neuroprotective; antibacterial; virucide; fungicide; ophthalmological;  
CC and vulnary. The polynucleotides and polypeptides can be used to  
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used  
CC in diagnosing a pathological condition or susceptibility to a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms or  
CC cancer of the breast or liver, cardiovascular disorders, cerebrovascular  
CC disorders, angiogenesis, nervous system disorders, infections caused by  
CC bacteria, viruses and fungi and ocular disorders. The polypeptides can  
CC also be used to aid wound healing and epithelial cell proliferation, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, for supporting cell culture of primary tissues, to  
CC regenerate tissues and in chemotaxis. The polypeptides can also be used  
CC as a food additive or preservative to increase or decrease storage  
CC capabilities. AAC59440 to AAC59448 and AAB34091 represent sequences used  
CC in the exemplification of the present invention.  
XX  
SQ Sequence 2267 BP; 741 A; 416 C; 490 G; 617 T; 3 other;

Query Match

Best Local Similarity 56.5%; Score 2092; DB 21; Length 2267;

Matches 2192; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1496 GAAGGCTTTGGGCACAGACCACACAGGATCTATGGCAGCAAAAATGAACGAGAG 1555

Db 32 GAAGGCTTTGGGCACAGACCACACAGGATCTATGGCAGCAAAAATGAACGAGAG 91

Qy 1556 ATTTTGGAGAGACCCAGCTTGCATATGCAAGGAGACCCCTCACAGCCAGGGGTA 1615

Db 92 ATTTTGGAGAGACCCAGCTTGCATATGCAAGGAGACCCCTCACAGCCAGGGGTA 151

Qy 1616 GAAGATGCCCCCTATTGCTAAGGTGGTGTCTCGCTGCAAGTATGGAACCAAGGCTCC 1675

Db 152 GAAGATGCCCCCTATTGCTAAGGTGGTGTCTCGCTGCAAGTATGGAACCAAGGCTCC 211

Qy 1676 TCTCAGCAGGAGGAAGGAGCAAGCCAGCTGAAACCAAGCTGAGGATCGCTTGCCA 1735

Db 212 TCTCAGCAGGAGGAAGGAGCAAGCCAGCTGAAACCAAGCTGAGGATCGCTTGCCA 271



gene therapy; cancer; ss.  
Homo sapiens.  
Key  
CDS  
Location/Qualifiers  
474..2276  
/\*tag=a  
/transl\_except= "(pos: 825..827, aa: Val)"  
/transl\_except= "(pos: 1479..1481, aa: Asp)"  
/transl\_except= "(pos: 1627..1629, aa: Lys)"  
/transl\_except= "(pos: 1860..1861, aa: Ser)"  
/transl\_except= "(pos: 1968..1970, aa: Ser)"  
/transl\_except= "(pos: 2031..2033, aa: Ile)"  
/transl\_except= "(pos: 2088..2090, aa: Ser)"  
/transl\_except= "(pos: 2106..2108, aa: Ser)"  
/product= "EPLIN (epithelial protein lost in neoplasm)"  
WO200118019-A1.  
XX  
XX  
PD 15-MAR-2001.  
XX  
PF 08-SEP-2000; 2000WO-US24689.  
XX  
PR 08-SEP-1999; 99US-0153024.  
XX  
PA (REGC ) UNIV CALIFORNIA.  
XX  
Chang DD, Maul RS;  
WPI: 2001-244555/25.  
DR P-PSDB, AAB67700.  
XX  
XX New tumor suppressor protein EPLIN, useful as a marker for diagnostic,  
PT prognostic and therapeutic applications over the course of cell  
PT proliferative disorders associated with EPLIN -  
XX  
PS Claim 4; Page 42-43; 59pp; English.  
XX  
CC The present sequence encodes a human EPLIN (epithelial protein lost in  
CC neoplasm)-alpha isoform. The specification also describes EPLIN-beta.  
CC EPLIN is a tumor suppressor protein, whose expression is altered in  
CC multiple common human tumor types. EPLIN nucleic acids and proteins are  
CC used in screening assays to detect molecules that specifically bind to  
CC EPLIN nucleic acids, proteins or derivatives and thus have potential use  
CC as agonist or antagonist of EPLIN, in particular molecules that affect  
CC cell proliferation. Thus the assays are useful for screening molecules  
CC with potential utility as anticancer drugs or lead compounds for drug  
CC development. EPLIN nucleic acids, proteins are useful for detecting a  
CC cell proliferative disorder in a subject. EPLIN polynucleotides are  
CC useful in gene therapy techniques. EPLIN is useful as a marker that  
CC can be diagnostically, prognostically and therapeutically used over  
CC the course of a cell proliferative disorder associated with EPLIN.  
XX  
SQ Sequence 3543 BP; 1152 A; 711 C; 771 G; 909 T; 0 other;

Query Match 55.1%; Score 2041; DB 22; Length 3543;  
Best Local Similarity 99.4%; Pred No. 0;  
Matches 3371; Conservative 0; Mismatches 11; Indels 10; Gaps 3;  
QY 302 AACACCGAAATCTCTCCAGACCTTTAGAAAGGGGACCTGACTGTGTTAAAGAGAG 361  
DB 159 AACACCGAAATCTCTCCAGACCTTTAGAAAGGGGACCTGACTGTGTTAAAGAGAG 218  
QY 362 TGGGAGACCCAGGCTGGGAGCAGAGTCTCACAGACTCTCTACGGAAACAGCAGCACT 421  
DB 219 TGGGAGACCCAGGCTGGGAGCAGAGTCTCACAGACTCTCTACGGAAACAGCAGCACT 278  
QY 422 GAGATTAGGCACAGACAGACCATCTCTGCTGAAGTGACAGCCAGCTGCTTCTTGG 481  
DB 279 GAGATTAGGCACAGACAGACCATCTCTGCTGAAGTGACAGCCAGCTGCTTCTTGG 338  
QY 482 GCCAAAGTGTACCAAGAGAAACAATCCACCCAGACTAGACTCAGTCTCAGCTCCTGAA 541

DB 339 GCCAAAGCTGACCAAGAGAAACAAATCCACCAGATCTTAGACTCAGGTCACTCTGTAA 398  
QY 542 GCCCTCGTTCCAGGTGGATATCCACATCAAGGACGGTGAGGATCTTAAAGACCACTCA 601  
DB 399 GCCCTCGTTCCAGGTGGATATCCACATCAAGGACGGTGAGGATCTTAAAGACCACTCA 458  
QY 602 ACAGAAAGTAAAGAAATGGAATTTGTTAGGAAATCCAGGATGAAGTAGAAAAATCA 661  
DB 459 ACAGAAAGTAAAGAAATGGAATTTGTTAGGAAATCCAGGATGAAGTAGAAAAATCA 518  
QY 662 GAAATCAGTCAAAACACAGATGCTTCGGGCAAAATAGAGAAATATATATGTTCCGCTGAAC 721  
DB 519 GAAATCAGTCAAAACACAGATGCTTCGGGCAAAATAGAGAAATATATATGTTCCGCTGAAC 578  
QY 722 AGCTTTAAGATGATGTTTGAGAAAGTGAAACCACTCAAACTAAGATTTCTCGGGGCCAA 781  
DB 579 AGCTTTAAGATGATGTTTGAGAAAGTGAAACCACTCAAACTAAGATTTCTCGGGGCCAA 638  
QY 782 AGCCGAAGTGCAGTGAAGGAGGAGATCTCTGAAACAGCTATTCTCTAGATGACCTGGAA 841  
DB 639 AGCCGAAGTGCAGTGAAGGAGGAGATCTCTGAAACAGCTATTCTCTAGATGACCTGGAA 698  
QY 842 ATAGGCCAGGTGAGTTGTCATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGA 901  
DB 699 ATAGGCCAGGTGAGTTGTCATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGA 758  
QY 902 AATCTGGAATTTCCAGCGCTCTCAGAAACCTCTATAAAGGATCGAATGCCCAAGTACAG 961  
DB 759 AATCTGGAATTTCCAGCGCTCTCAGAAACCTCTATAAAGGATCGAATGCCCAAGTACAG 818  
QY 962 GCAGTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGGT 1021  
DB 819 GCAGTGTGTCCAAACAAAGCAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGGT 878  
QY 1022 GCGGAAATCAAAATTTCAATAAATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTCAGGTC 1081  
DB 879 GCGGAAATCAAAATTTCAATAAATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTCAGGTC 938  
QY 1082 TGCATCACCCATCAGGAGGAGGAGAAATTTCTGCAAAATGAGAAATAGCTGGAGTCCGT 1141  
DB 939 TGCATCACCCATCAGGAGGAGGAGAAATTTCTGCAAAATGAGAAATAGCTGGAGTCCGT 998  
QY 1142 TCCACCCCTGCGGAGGAGTACTCCC---GTCACTCCAGGTTAAGAGTGAAGTTCAACAG 1198  
DB 999 TCCACCCCTGCGGAGGAGTACTCCCAGGTTAAGAGTGAAGTTCAACAG 1058  
QY 1199 CTTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGTCTTTCTGAAAGTTCT 1258  
DB 1059 CTTGTCCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCCAGTCTTTCTGAAAGTTCT 1118  
QY 1259 CTTCCCAAGCAATGAAGAGTTTTCAGGCACTGCAAGAGAGACCTCGGTGGAATGTGAG 1318  
DB 1119 CTTCCCAAGCAATGAAGAGTTTTCAGGCACTGCAAGAGAGACCTCGGTGGAATGTGAG 1178  
QY 1319 AAGACAGTCTATCAATGGAGCGTCTCTTGGCCAAACAGCAGGTGTTTACATCAGCTGC 1378  
DB 1179 AAGACAGTCTATCAATGGAGCGTCTCTTGGCCAAACAGCAGGTGTTTACATCAGCTGC 1238  
QY 1379 TTCGGTTGCTCTTATGCAACCAAACTCAGTCTAGGAACATATGCATCTTTACATGGA 1438  
DB 1239 TTCGGTTGCTCTTATGCAACCAAACTCAGTCTAGGAACATATGCATCTTTACATGGA 1298  
QY 1439 AGAATCTATTGTAAGCTCACTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGA 1498  
DB 1299 AGAATCTATTGTAAGCTCACTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGA 1358  
QY 1499 GGCTTTGGGCACAGACCAACAGAGATCTATGGGCAAGCAAAATGAAACGAGAGATT 1558  
DB 1359 GGCTTTGGGCACAGACCAACAGAGATCTATGGGCAAGCAAAATGAAACGAGAGATT 1418  
QY 1559 TTGGAGAGACCCAGCTTGCAGAAATGCAAGGAGACCCCTCAGAGCCAGGGGTAGAA 1618  
DB 1419 TTGGAGAGACCCAGCTTGCAGAAATGCAAGGAGACCCCTCAGAGCCAGGGGTAGAA 1478







AC AAH18125;  
XX  
XX 26-JUN-2001 (first entry)  
DT  
XX Human cDNA sequence SEQ ID NO:17991.  
DE  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
XX Homo sapiens.  
OS  
XX EP1074617-A2.  
XX  
XX PD 07-FEB-2001.  
XX  
XX PF 28-JUL-2000; 2000EP-0116126.  
XX  
XX PR 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-018776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 17991; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
XX Sequence 2207 BP; 692 A; 438 C; 512 G; 565 T; 0 other;  
SQ

Query Match 41.0%; Score 1519; DB 22; Length 2207;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1569; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
1001 AATGAGCTGAAGCCAGTGGTGGGAAATCAAAATTCATAAAATCGAGCAAAAGAGAAAT 1060  
638 AATGAGCTGAAGCCAGTGGTGGGAAATCAAAATTCATAAAATCGAGCAAAAGAGAAAT 697  
1061 GTGCCCCAGTCTCTGAGTCTGATCACCATCCCATCAGGAAGGGGAAAGATTTCTGCAAT 1120  
698 GTGCCCCAGTCTCTGAGTCTGATCACCATCCCATCAGGAAGGGGAAAGATTTCTGCAAT 757

QY 1121 GAGAAATAGCCTGGCAGTCCGTTTCCACCCCTGCGAAGATGACTCCCGTACTCCAGGTT 1180  
Db |||||  
758 GAGAAATAGCCTGGCAGTCCGTTTCCACCCCTGCGAAGATGACTCCCGTACTCCAGGTT 817  
QY |||||  
1181 AAGAGTGAAGTTCAACAGCCTGTCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 1240  
Db |||||  
818 AAGAGTGAAGTTCAACAGCCTGTCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 877  
QY |||||  
1241 AGTCTTTCTGAAAGTTCTCTCTCCCAAGCAATGAAGATTTTCAGGCACCTGCAAGAGAG 1300  
Db |||||  
878 AGTCTTTCTGAAAGTTCTCTCTCCCAAGCAATGAAGATTTTCAGGCACCTGCAAGAGAG 937  
QY |||||  
1301 ACTGCGTGAATGTGAGAGCAGTCTATCCAAATGGAGCGTCTCTTGGCCAAACCCAGCAG 1360  
Db |||||  
938 ACTGCGTGAATGTGAGAGCAGTCTATCCAAATGGAGCGTCTCTTGGCCAAACCCAGCAG 997  
QY |||||  
1361 GTGTTTTCAGATCAGCTGCTTCCGTTCTCTATTTGCAACAACTCAGTCTAGGAACA 1420  
Db |||||  
998 GTGTTTTCAGATCAGCTGCTTCCGTTCTCTATTTGCAACAACTCAGTCTAGGAACA 1057  
QY |||||  
1421 TATGCATCTTTACATGGAAGAATCTATTGTAAGCCTCACTTCAATCAACTCTTTAAATCT 1480  
Db |||||  
1058 TATGCATCTTTACATGGAAGAATCTATTGTAAGCCTCACTTCAATCAACTCTTTAAATCT 1117  
QY |||||  
1481 AAGGGCAACTATGATGAAGGCTTTGGGCACACACACACACAGGATCTATGGCAAGCACA 1540  
Db |||||  
1118 AAGGGCAACTATGATGAAGGCTTTGGGCACACACACACACAGGATCTATGGCAAGCACA 1177  
QY |||||  
1541 AATGAAAACGAAGAGATTTTGGAGAGACAGCCAGCTTGAAGTGAAGGAGAGACCCCT 1600  
Db |||||  
1178 AATGAAAACGAAGAGATTTTGGAGAGACAGCCAGCTTGAAGTGAAGGAGAGACCCCT 1237  
QY |||||  
1601 CACAGCCAGGGGTAGAAGATGCCCCCTATTGTAAGTGGGTGCTCTGGCTGCAAGTATG 1660  
Db |||||  
1238 CACAGCCAGGGGTAGAAGATGCCCCCTATTGTAAGTGGGTGCTCTGGCTGCAAGTATG 1297  
QY |||||  
1661 GAAGCAAGAGGCTCTCTCTCAGCAGAGAGAGGAGACAGCAGCTGAAACCAAGAGCTG 1720  
Db |||||  
1298 GAAGCAAGAGGCTCTCTCTCAGCAGAGAGAGGAGACAGCAGCTGAAACCAAGAGCTG 1357  
QY |||||  
1721 AGGATCGCTCGCCACCCCTCACTGAACTTGAAGTTCAGGAAGTGCCTTGGAGAAAGG 1780  
Db |||||  
1358 AGGATCGCTCGCCACCCCTCACTGAACTTGAAGTTCAGGAAGTGCCTTGGAGAAAGG 1417  
QY |||||  
1781 ATCAAAATGTCAAGCCCAAAATGGCTCTCTGAGACGAAATCAGCAAGCCCAAGTTCCT 1840  
Db |||||  
1418 ATCAAAATGTCAAGCCCAAAATGGCTCTCTGAGACGAAATCAGCAAGCCCAAGTTCCT 1477  
QY |||||  
1841 GAGGATGTCGATCTAGATCTCAAGAAAGCTAAGACGATCTCTTCACTGAAGGAAAGAGC 1900  
Db |||||  
1478 GAGGATGTCGATCTAGATCTCAAGAAAGCTAAGACGATCTCTTCACTGAAGGAAAGAGC 1537  
QY |||||  
1901 CGCCCATTCAGTGTAGCAGCTTCAATTTCAAGACACCTCTGTCAGAGAGCCCAAAACTGTG 1960  
Db |||||  
1538 CGCCCATTCAGTGTAGCAGCTTCAATTTCAAGACACCTCTGTCAGAGAGCCCAAAACTGTG 1597  
QY |||||  
1961 TCCCCACTATCAGGAAAGGCTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGGTGGA 2020  
Db |||||  
1598 TCCCCACTATCAGGAAAGGCTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGGTGGA 1657  
QY |||||  
2021 AGAGTTTCAGAAAGGAAACAAAGTGGAAAAATGCAAGGCTTCTTAAGAGAAATGGGAATGTG 2080  
Db |||||  
1658 AGAGTTTCAGAAAGGAAACAAAGTGGAAAAATGCAAGGCTTCTTAAGAGAAATGGGAATGTG 1717  
QY |||||  
2081 GGAAGAAACAACTGGCAACCAACCAAGATCTTAAGGAGAGACAGGGAAGAGAAATAGGAA 2140  
Db |||||  
1718 GGAAGAAACAACTGGCAACCAACCAAGATCTTAAGGAGAGACAGGGAAGAGAAATAGGAA 1777  
QY |||||  
2141 GGTCAATAGTTTGGAGATGGAGATGAGATCTTTAGAAATGTTGTCAGACCTCCGATGAA 2200  
Db |||||  
1778 GGTCAATAGTTTGGAGATGGAGATGAGATCTTTAGAAATGTTGTCAGACCTCCGATGAA 1837  
QY |||||  
2201 GATGATAACAGCTTCTCTCAACCAACCAATCTCCCAAGAACCCCAAGTCTCTGAAATGGTGC 2260

Db 1838 GATGATAACAGCTTCTCAACAAACAATCTCCACAGAACCCAAAGTCTCTGAATTGGTCG 1897  
QY 2261 AGTTTGTAGACACACCTTCTCTGAAGAAATCTACTACTCAGAAATCCAGGAT 2320  
Db 1898 AGTTTGTAGACACACCTTCTCTGAAGAAATCTACTACTCAGAAATCCAGGAT 1957  
QY 2321 GTGGAATCTCGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAGAGA 2380  
Db 1958 GTGGAATCTCGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAGAGA 2017  
QY 2381 AATCGGTATTATGATGAGGATGAGTGAAGAGTGAACAAATTCGAATGATGCTGGGCTT 2440  
Db 2018 AATCGGTATTATGATGAGGATGAGTGAAGAGTGAACAAATTCGAATGATGCTGGGCTT 2077  
QY 2441 AATTCATGTTAGTGTAGCGAGCAGTCCCTTTGTCAAAATGTGATGACATAAGCAG 2500  
Db 2078 AATTCATGTTAGTGTAGCGAGCAGTCCCTTTGTCAAAATGTGATGACATAAGCAG 2137  
QY 2501 GTATCCCGACATGAATGTAATTTACTTGGAGTAACCTTGGAAAGAAATTCCTTCTTAA 2560  
Db 2138 GTATCCCGACATGAATGTAATTTACTTGGAGTAACCTTGGAAAGAAATTCCTTCTTAA 2197  
QY 2561 AATCAAAAAC 2570  
Db 2198 AATCAAAAAC 2207

RESULT 11  
AA159955  
ID AA159955 standard; cDNA; 1713 BP.  
XX AA159955;  
AC AA159955;  
DT 22-OCT-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 3944.  
DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
leukaemia; ss.  
XX Homo sapiens.  
OS  
XX WO200153312-A1.  
PN  
XX 26-JUL-2001.  
PD  
XX 26-DEC-2000; 2000WO-US34263.  
PF  
XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 09-JUL-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0820312.  
PR 03-AUG-2000; 2000US-0853450.  
PR 14-SEP-2000; 2000US-0662191.  
PR 19-OCT-2000; 2000US-0693036.  
PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR P-PSDB; AAW40799.  
XX Novel nucleic acids and polypeptides, useful for treating disorders  
PT such as central nervous system injuries.

XX PS Claim 1; SEQ ID NO 3944; 10078pp; English.  
XX The invention relates to human nucleic acids (AA157798-AA161369) and  
CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemias and  
CC C.N.S disorders.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification.  
XX  
SQ Sequence 1713 BP; 549 A; 401 C; 417 G; 346 T; 0 other;  
Query Match 35.4%; Score 1310; DB 22; Length 1713;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1510; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 22 TAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGAGCGCGCGGACCTGTGACAGGGCTGGT 81  
Db 6 TAGCAGCTTGGTCGCGACAGGTGCGCTAGGTAGAGCGCGCGGACCTGTGACAGGGCTGGT 65  
QY 82 AGCAGCGCAGAGGAAGCGCGCTTTTAGCCAGGTATTTAGTGTCTGTAGACAAGATGA 141  
Db 66 AGCAGCGCAGAGGAAGCGCGCTTTTAGCCAGGTATTTAGTGTCTGTAGACAAGATGA 125  
QY 142 ATCATCTCCATTTAATAGACGGCAATGGACCTCCTATCATTTGAGGGTAAACAGCCAAAGA 201  
Db 126 ATCATCTCCATTTAATAGACGGCAATGGACCTCCTATCATTTGAGGGTAAACAGCCAAAGA 185  
QY 202 ACTTTCTCTTGCAACAAGCAAGTCACTCGCTATTGTGGAATATTCTCCAAGTACCA 261  
Db 186 ACTTTCTCTTGCAACAAGCAAGTCACTCGCTATTGTGGAATATTCTCCAAGTACCA 245  
QY 262 GAAAGCAGCTGAGAAACAACATCGAGAGAGAGTAACACCGAAAAATCTCTCCCA 321  
Db 246 GAAAGCAGCTGAGAAACAACATCGAGAGAGAGTAACACCGAAAAATCTCTCCCA 305  
QY 322 GCACCTTTAGAAAGGGGACCTTGACTGTGTTAAGAGAGAGTGGGAGAACCCAGGCTGGG 381  
Db 306 GCACCTTTAGAAAGGGGACCTTGACTGTGTTAAGAGAGAGTGGGAGAACCCAGGCTGGG 365  
QY 382 AGCAGAGTCTCACACAGACTCTCTACGGAACAGCAGCAGTCTAGATTAGGCACAGAGCAGA 441  
Db 366 AGCAGAGTCTCACACAGACTCTCTACGGAACAGCAGCAGTCTAGATTAGGCACAGAGCAGA 425  
QY 442 CCATCTCTCTGCTGAAGTGAACGCCAGCTGCTTTCTGGAGCCAAAGCTGACCAAGAGA 501  
Db 426 CCATCTCTCTGCTGAAGTGAACGCCAGCTGCTTTCTGGAGCCAAAGCTGACCAAGAGA 485  
QY 502 ACAAATCCACCCCGAGATCTAGACTCAGTCACTCTCTGAAAGCCCTGTTCAAGGTCGATA 561  
Db 486 ACAAATCCACCCCGAGATCTAGACTCAGGCGACCTCTCTGAAAGCCCTGTTCAAGGTCGATA 545  
QY 562 TCCCCACATCAAGGACGCTGAGGATCTTAAAGACCACTCAACAGAAAGTAAAAAATGGA 621  
Db 546 TCCCCACATCAAGGACGCTGAGGATCTTAAAGACCACTCAACAGAAAGTAAAAAATGGA 605  
QY 622 AAATTGTCTAGGAGAAATCCAGGATGAAGTAGAAAAATCAGAAATCAGTGTGAAACACAGA 681  
Db 606 AAATTGTCTAGGAGAAATCCAGGATGAAGTAGAAAAATCAGATCAGTGTGAAACACAGA 665  
QY 682 TGCTTCGGGCAAAATAGAGAAATATATGTTCCGCTGAACAGGCTTAAAGATGATTTGA 741  
Db 666 TGCTTCGGGCAAAATAGAGAAATATATGTTCCGCTGAACAGGCTTAAAGATGATTTGA 725

QY 742 GAAAGGTGAACCAACTCAAAGTATCTCCGGGCCCAAGCCGAAGTCAAGTGGAG 801  
DB |||||||  
QY 726 GAAAGGTGAACCAACTCAAAGTATCTCCGGGCCCAAGCCGAAGTCAAGTGGAG 785  
DB |||||||  
QY 802 CAAGATCTCTCAAAACAGCTATTCTCTAGATGACCTCGAAATAGGCCAGCTGAGTGTG 861  
DB |||||||  
QY 786 GAAGATCTCTGAAACAGCTATTCTCTAGATGACCTCGAAATAGGCCAGCTGAGTGTG 845  
DB |||||||  
QY 862 ATCTTCTACATTTGACTCGGAGAAATAGAGTAGACGAAATCTGGAATCTCCAGCCT 921  
DB |||||||  
QY 846 ATCTTCTACATTTGACTCGGAGAAATAGAGTAGACGAAATCTGGAATCTCCAGCCT 905  
DB |||||||  
QY 922 CTCAGAAACCTCTATAAGAGTCAATGGCCAAGTACCGGAGCTGTGTCCAAACAAAG 981  
DB |||||||  
QY 906 CTCAGAAACCTCTATAAGAGTCAATGGCCAAGTACCGGAGCTGTGTCCAAACAAAG 965  
DB |||||||  
QY 982 CAGCTCAACCAACTATACAAATGAGCTGAAAGCAGTGTGGGAAATCAAAATTCATAA 1041  
DB |||||||  
QY 966 CAGCTCAACCAACTATACAAATGAGCTGAAAGCAGTGTGGGAAATCAAAATTCATAA 1025  
DB |||||||  
QY 1042 AATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTGAGTCTGCATCACCCATCAGGAAG 1101  
DB |||||||  
QY 1026 AATGGAGCAAAAGGAGAAATGTGCCCCAGGTCTGAGTCTGCATCACCCATCAGGAAG 1085  
DB |||||||  
QY 1102 GGAAGAGATTTCTGCAATGAGAAATAGCCTGGCAGTCCGTTCACCCCTGCCGAAGTGA 1161  
DB |||||||  
QY 1086 GGAAGAGATTTCTGCAATGAGAAATAGCCTGGCAGTCCGTTCACCCCTGCCGAAGTGA 1145  
DB |||||||  
QY 1162 CTCCTGAGTCTCCAGGTAAAGAGTGAAGTCAAGAGCTGTCCATGCCAAGCCACTAAG 1221  
DB |||||||  
QY 1146 CTCCTGAGTCTCCAGGTAAAGAGTGAAGTCAAGAGCTGTCCATGCCAAGCCACTAAG 1205  
DB |||||||  
QY 1222 TCAGATTTCCAGAGCTCCAGTCTTTCTGAAAGTCTCTCCCAAGCAATGAAGAGTT 1281  
DB |||||||  
QY 1206 TCAGATTTCCAGAGCTCCAGTCTTTCTGAAAGTCTCTCCCAAGCAATGAAGAGTT 1265  
DB |||||||  
QY 1282 TCAGGACCTGCAAGAGAGACCTGCGTGAATGTGAGAGACAGTCTATCCCAATGGAGCG 1341  
DB |||||||  
QY 1266 TCAGGACCTGCAAGAGAGACCTGCGTGAATGTGAGAGACAGTCTATCCCAATGGAGCG 1325  
DB |||||||  
QY 1342 TCTCTGGCCAAACAGCAGGTGTTTCAATCAGCTGCTTCGTTGCTCTATTGGCAACAA 1401  
DB |||||||  
QY 1326 TCTCTGGCCAAACAGCAGGTGTTTCAATCAGCTGCTTCGTTGCTCTATTGGCAACAA 1385  
DB |||||||  
QY 1402 CAAACTCAGCTAGGAACATATGATCTTTTACATGGAAGATCTATTAAGCCCTCACTT 1461  
DB |||||||  
QY 1386 CAAACTCAGCTAGGAACATATGATCTTTTACATGGAAGATCTATTAAGCCCTCACTT 1445  
DB |||||||  
QY 1462 CAACTCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACACACAA 1521  
DB |||||||  
QY 1446 CAACTCAACTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACACACAA 1505  
DB |||||||  
QY 1522 GGATCTATGGCAA 1535  
DB |||||||  
QY 1506 GGATCTATGGCAA 1519  
DB |||||||

RESULT 12  
ABI98973  
ID ABI98973 standard; cDNA; 1754 BP.  
AC ABI98973;  
XX  
AC ABI98973;  
XX  
DT 21-FEB-2002 (first entry)  
XX  
DE Human cancer suppressor protein PP624 encoding cDNA.  
XX  
KW Human; cancer suppressor; disease; cancer; ss.  
XX  
DS Homo sapiens.  
XX  
FH Key Location/Qualifiers

CDS 444..1349  
FT FT /\*tag= a  
FT FT /transl\_except= (pos:104..106,aa:Xaa)  
FT FT /transl\_except= (pos:561..563,aa:Xaa)  
FT FT /transl\_except= (pos:570..572,aa:Xaa)  
FT FT /product= "PP624"  
FT FT /note= "Xaa = unknown"  
XX  
PN CN1313297-A.  
XX  
XX 19-SEP-2001.  
XX  
PF 09-MAR-2000; 2000CN-0111948.  
XX  
PR 09-MAR-2000; 2000CN-0111948.  
XX  
XX (SHAN-) SHANGHAI INST ONCOLOGY.  
XX  
XX Gu J, Yang S;  
XX  
XX WPI; 2002-042185/D6.  
XX P-PSDB; ABB56420.  
XX  
XX Human protein able to suppress growth of cancer cells and its coding  
XX sequence -  
XX  
XX Claim 5; Page 21-22 Disclosure; 37pp; Chinese.  
XX  
XX The invention relates to novel human proteins (ABB56417-ABB56425) with  
XX cancer suppressing function, the encoding polynucleotides  
XX (ABI98970-ABI98978), the process for preparing the polypeptide, the  
XX application of the polypeptide in treating diseases such as cancer, the  
XX antagonist of the polypeptide and its medical function and the  
XX application of the polynucleotide.  
XX  
SQ Sequence 1754 BP; 586 A; 377 C; 426 G; 361 T; 4 other;

Query Match 35.0%; Score 1296; DB 24; Length 1754;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1736; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 829 AGATGACCTGGAATAGGCCAGGTGATGTTGATCTTCTACATTTGACCTCGGAGAAAA 888  
DB 1 AGATGACCTGGAATAGGCCAGGTGATGTTGATCTTCTACATTTGACCTCGGAGAAAA 60  
QY 889 TGAGGTAGACGAAATCTGGAATTTCCAGCTCTCAGAAACCTCTATAAGAGTCAAT 948  
DB 61 TGAGGTAGACGAAATCTGGAATTTCCAGCTCTCAGAAACCTCTATAAGAGTCAAT 120  
QY 949 GGCCAGTACCCAGCAGCTGTGTCCAAACAAAGCAGCTCAACCACTATACAAAATGAGCT 1008  
DB 121 GGCCAGTACCCAGCAGCTGTGTCCAAACAAAGCAGCTCAACCACTATACAAAATGAGCT 180  
QY 1009 GAAAGCCAGTGTGGGCAAAATCAAAATTTCAATAATGGAGCAAAAGAGAAATGTGCCCC 1068  
DB 181 GAAAGCCAGTGTGGGCAAAATCAAAATTTCAATAATGGAGCAAAAGAGAAATGTGCCCC 240  
QY 1069 AGTCTCTGAGGTCTGCATCAACCTTCCAGAGGGAAGAGATTTCTGCAAAATGAGAAATAG 1128  
DB 241 AGTCTCTGAGGTCTGCATCAACCTTCCAGAGGGAAGAGATTTCTGCAAAATGAGAAATAG 300  
QY 1129 CCTGGCAGTCCGTTCCAGCTTCCAGAGGGAAGAGATTTCTGCAAAATGAGAAATAG 1188  
DB 301 CCTGGCAGTCCGTTCCAGCTTCCAGAGGGAAGAGATTTCTGCAAAATGAGAAATAG 360  
QY 1189 GGTTCAAAGCAGCTGTCCATCCCAAGCCACTAAGTCCAGAT-TCCAGAGCTTCCAGTCTTT 1247  
DB 361 GGTTCAAAGCAGCTGTCCATCCCAAGCCACTAAGTCCAGAT-TCCAGAGCTTCCAGTCTTT 420  
QY 1248 CTGAAAGTTCTCTCCCAAGCAATGAAGAGTTTTCAGGCACTTTCAGGCACTTTCAGGCACTTTCAGG 1307  
DB 421 CTGAAAGTTCTCTCCCAAGCAATGAAGAGTTTTCAGGCACTTTCAGGCACTTTCAGGCACTTTCAGG 480



CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,  
CC carbohydrates, vitamins, minerals, cofactors and other nutritional  
CC components. Oligonucleotides AAC93470 - AAC93478 and peptide AA951926 are  
CC used in the isolation and characterisation of the proteins and  
CC polynucleotides of the invention.  
XX  
SQ Sequence 2158 BP; 714 A; 383 C; 458 G; 603 T; 0 other;

```

Query Match          34.3%; Score 1269; DB 21; Length 2158;
Best Local Similarity 99.5%; Pred. NO. 0;
Matches 2129; Conservative 0; Mismatches 4; Indels 6

```

QY	1553	GAGATTTTGGAGAGACCCAGCCAGCTTGCATAATGCAAGGGAGACCCCTCACAGCCCAAGG	1612
Db	1	GAGATTTTGGAGAGACCCAGCCAGCTTGCATAATGCAAGGGAGACCCCTCACAGCCCAAGG	60
QY	1613	GTAGAAGATGCCCTTATTGCTAAGTGGGTGTCTGGCTGCAAGATATGGAAGCCCAAGGCC	1672
Db	61	GTAGAAGATGCCCTTATTGCTAAGTGGGTGTCTGGCTGCAAGATATGGAAGCCCAAGGCC	120
QY	1673	TCCTCTCAGCAGGAGGAAGACACAGCAGCTGAACCAAGAGCTGAGGATCGCCTGG	1732
Db	121	TCCTCTCAGCAGGAGGAAGACACAGCAGCTGAACCAAGAGCTGAGGATCGCCTGG	180
QY	1733	CCACCCCCACTGAACTTTGGAAGTTTCAGAAAGTGCCTTGGAGGAAGGGATCAAAATGTCA	1792
Db	181	CCACCCCCACTGAACTTTGGAAGTTTCAGAAAGTGCCTTGGAGGAAGGGATCAAAATGTCA	240
QY	1793	AAGCCCAATGGCTCTCTGAAGACGAAATCAGCAAGCCCGAAGTTCTCTGAGGATGTTCGAT	1852
Db	241	AAGCCCAATGGCTCTCTGAAGACGAAATCAGCAAGCCCGAAGTTCTCTGAGGATGTTCGAT	300
QY	1853	CTAGATCTGAAGAAGCTAAGACGATCTTCTTCTACTGAAGAAAGACGCCCAATTCCT	1912
Db	301	CTAGATCTGAAGAAGCTAAGACGATCTTCTTCTACTGAAGAAAGACGCCCAATTCCT	360
QY	1913	GTAGCAGCTTCATTTCAAAGCACCTCTGCAAGAGCCCAAAAACCTGTGCCCACTATC	1972
Db	361	GTAGCAGCTTCATTTCAAAGCACCTCTGCAAGAGCCCAAAAACCTGTGCCCACTATC	420
QY	1973	AGGAAAGGCTGGAGCATGTTCAGAGCAGAGTGAAGAGTCTGTGGTGGAAAGAGTTGCAGAA	2032
Db	421	AGGAAAGGCTGGAGCATGTTCAGAGCAGAGTGAAGAGTCTGTGGTGGAAAGAGTTGCAGAA	480
QY	2033	AGGAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAAATGGGAATGGGAAAAACAACC	2092
Db	481	AGGAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAAATGGGAATGGGAAAAACAACC	540
QY	2093	TGGCAAAACAAAGATCTTAAAGAGAGACAGGGAAGAGAGTGAAGAGGTCAATAGTTTG	2152
Db	541	TGGCAAAACAAAGATCTTAAAGAGAGACAGGGAAGAGAGTGAAGAGGTCAATAGTTTG	600
QY	2153	GAGATGGAGAAATCAGAAATCTTGTAGAAAAATGGTGCAGACTCCGATGAAGATCATACAGC	2212
Db	601	GAGATGGAGAAATCAGAAATCTTGTAGAAAAATGGTGCAGACTCCGATGAAGATCATACAGC	660
QY	2213	TTCTCTAAACAAATCTCCAC--AAGAACCCAAAGTCTCTGAATTTGGTCGAGTTTG-TA	2269
Db	661	TTCTCTAAACAAATCTCCACCAAGGAACCCAAAGTCTCTGAATTTGGTCGAGTTTGTTA	720
QY	2270	GACAAACACCTTT-GCTGAAGAA-TTCACCTACTCAGAAATCCAGGATGTGGAAC	2327
Db	721	GACAAACACCTTTGGCTGAAGAAATTTTCACTACTCAGAAATCCAGGATGTGGAAC	780
QY	2328	TGTGGAGGGAAGTGGTCAAGAGCTCTCTGTGGAAGACAGATAAGAGAATCGGT	2387

Dh		781	TCTGGGAGGAGAGTGGTCAAAGAGCTCTCTGTGGAAGAACAAGATAAAAGAGAAATCGGT	840
Qy		2388	ATTATGATGAGGATGAGGATGAAGAGTGACAAATTCGAATGATGCTGGGCCTTTAAATTC	2447
Dh		841	ATTATGATGAGGATGAGGATGAAGAGTGACAAATTCGAATGATGCTGGGCCTTTAAATTC	900
Qy		2448	TGTTAGTGTTTAGCGAGCCACTGCCCTTTGCAAAAATGTCATGCACATAAGCAGGTATCCC	2507
Dh		901	TGTTAGTGTTTAGCGAGCCACTGCCCTTTGCAAAAATGTCATGCACATAAGCAGGTATCCC	960
Qy		2508	AGCATGAAATGTAATTTACTTGGAGTAACCTTTGGAAAAAGAAATTCCTCTCTTAAATTCAAA	2567
Dh		961	AGCATGAAATGTAATTTACTTGGAGTAACCTTTGGAAAAAGAAATTCCTCTCTTAAATTCAAA	1020
Qy		2568	AACAAAACAAAAAACAACAAAAACACATCTCTAAATCTAGAGATAACTTTACTTTAAAT	2627
Dh		1021	AACAAAACAAAAAACAACAAAAACACATCTCTAAATCTAGAGATAACTTTACTTTAAAT	1080
Qy		2628	CCTTCATTTTAGCAGTCATGATGATGATAAGTCTGTPAAGCCTTGTAACCTGGGGAATATT	2687
Dh		1081	CCTTCATTTTAGCAGTCATGATGATGATAAGTCTGTPAAGCCTTGTAACCTGGGGAATATT	1140
Qy		2688	CCACCTGATAATAGCCCCAGATTCTACTGTATTTCCCAAAGGCAATATTAAAGTAGATAGA	2747
Dh		1141	CCACCTGATAATAGCCCCAGATTCTACTGTATTTCCCAAAGGCAATATTAAAGTAGATAGA	1200
Qy		2748	TGATTAGTAGTATATGTTTACACACTATTTTGGAAATTAGAGAACATACAGAAAGGAATTA	2807
Dh		1201	TGATTAGTAGTATATGTTTACACACTATTTTGGAAATTAGAGAACATACAGAAAGGAATTA	1260
Qy		2808	GGGGCTTAAACATTAGGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATATTTTT	2867
Dh		1261	GGGGCTTAAACATTAGGACTGAATGCACCTTTAGTATAAAGGGCACAGTTTGTATATTTTT	1320
Qy		2868	AAATGAATACCAAATTAAATTTTGTAGTATTTTACCTGTTAAGAGATATTATTAGTCTTTAAA	2927
Dh		1321	AAATGAATACCAAATTAAATTTTGTAGTATTTTACCTGTTAAGAGATATTATTAGTCTTTAAA	1380
Qy		2928	TTTTTAGTGTAAATTTCTGCTGTGATATATATGAGGAATTTACTACTTTATGTCCTGC	2987
Dh		1381	TTTTTTAGTGTAAATTTCTGCTGTGATATATATGAGGAATTTACTACTTTATGTCCTGC	1440
Qy		2988	TCTCTAAACTACCTCGAACTCGACGCTCTGAGGTATAATACAAAGAGACACATTTTGA	3047
Dh		1441	TCTCTAAACTACCTCGAACTCGACGCTCTGAGGTATAATACAAAGAGACACATTTTGA	1500
Qy		3048	GGCAATCGAAAAACAACCTACACTCTTCGGTGTCTAGAGAGATCTGCTCTCCCAAAT	3107
Dh		1501	GGCAATCGAAAAACAACCTACACTCTTCGGTGTCTAGAGAGATCTGCTCTCCCAAAT	1560
Qy		3108	AAGCTTTTGATCTGCGCAGTGAATTTACTGACTCCAATGATGCTTTCTTTCTGGTG	3167
Dh		1561	AAGCTTTTGATCTGCGCAGTGAATTTACTGACTCCAATGATGCTTTCTTTCTGGTG	1620
Qy		3168	ATATCTGTGCTTCTCAATAATTACTGAAAGCTGCAATATTTTAGTAATACCTTCGGGATCA	3227
Dh		1621	ATATCTGTGCTTCTCAATAATTACTGAAAGCTGCAATATTTTAGTAATACCTTCGGGATCA	1679
Qy		3228	CTGTCCCCCATCTTCCGTGTTAGAGCAAGTGAAGAGTTTAAAGGAGGAGAAAGAA	3287
Dh		1680	CTGTCCCCCATCTTCCGTGTTAGAGCAAGTGAAGAGTTTAAAGGAGGAGAAAGAA	1739
Qy		3288	CTGTCTTTACACACTTTAGCTCAGACCTCTAAACCCCTGATTTTCCCTTATGATGTCCTCT	3347
Dh		1740	CTGTCTTTACACACTTTAGCTCAGACCTCTAAACCCCTGATTTTCCCTTATGATGTCCTCT	1799
Qy		3348	TTTTTGACACATAATTTTTTAAATACTTACTAGCTCTGAAAATATATTGATTTTTATCACAG	3407
Dh		1800	TTTTTGACACATAATTTTTTAAATACTTACTAGCTCTGAAAATATATTGATTTTTATCACAG	1859
Qy		3408	TATTTCTCAGGGTGAAATTAACCAACTATAGGCCCTTTTCTTGGGATGATTTCTTAGTCT	3467
Dh		1860	TATTTCTCAGGGTGAAATTAACCAACTATAGGCCCTTTTCTTGGGATGATTTCTTAGTCT	1919

QY 3468 TAAGTTTGGGACATTATAAAGTTGAGTACATTTGTTGACAGTTGATATTTCCAAAT 3527  
Db 1920 TAAGTTTGGGACATTATAAAGTTGAGTACATTTGTTGACAGTTGATATTTCCAAAT 1979  
QY 3528 TGTATGATGGGAGGAGGTGCTTTAAGCTGTAGGCTTTCTTTGTAAGTCAATTTATA 3587  
Db 1980 TGTATGATGGGAGGAGGTGCTTTAAGCTGTAGGCTTTCTTTGTAAGTCAATTTATA 2039  
QY 3588 GAGATTAGCTTTAATATTTTATAGAGATGATAAAGTCTGCTTTCTTTAGCTTTACCTA 3647  
Db 2040 GAGATTAGCTTTAATATTTTATAGAGATGATAAAGTCTGCTTTCTTTAGCTTTACCTA 2099  
QY 3648 GTCTGAACATTTTATTCATTAAGATTTTAATTTAAA 3686  
Db 2100 GTCTGAACATTTTATTCATTAAGATTTTAATTTAAA 2138

RESULT 14  
AAS18588  
ID AAS18588 standard; cDNA; 1567 BP.  
XX  
AC AAS18588;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE cDNA encoding human sterol regulatory element binding protein 3.  
DE  
XX Human; sterol regulatory element binding protein 3; hSREBP-3;  
KW hypothalamus; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 169..154  
FT CDS  
FT /\*tag=a  
FT /product="Sterol regulatory element binding protein 3,  
FT hSREBP-3"  
XX  
PN CN1309182-A.  
XX  
PD 22-AUG-2001.  
XX  
PF 17-FEB-2000; 2000CN-0111698.  
XX  
PR 17-FEB-2000; 2000CN-0111698.  
XX  
PA (NATU-) NATURAL HUMAN GENOME NANFANG RES CENT.  
XX  
PI Li Y, Xu S, Ren S;  
XX  
DR WPI; 2002-011822/02.  
DR P-PSDB; AAU10979.  
XX  
PT Cholesterol regulatory factor binding protein and its coding sequence -  
XX  
PS Claim 1; Page 22; 27pp; Chinese.  
XX  
CC The invention relates to a novel human sterol regulatory element binding  
CC protein 3 (hSREBP-3) expressed in human normal hypothalamic tissue and  
CC its coding sequence. Also described is the process for preparing the  
CC protein and nucleic acid sequence, and the method for detecting hSREBP-3  
CC nucleic acid sequence and polypeptides. The present sequence represents  
CC the coding sequence of human hSREBP-3 as described in the invention.  
XX  
SQ Sequence 1567 BP; 513 A; 326 C; 391 G; 337 T; 0 other;

Query Match 32.0%; Score 1186; DB 24; Length 1567;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1436; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1001 AATGAGCTGAAGCCAGTGTGGCGAAATCAAAATTCATAAAATGGAGCAAGGAAAT 1060  
|||||

Db 127 AATGAGCTGAAGCCAGTGTGGCGAAATCAAAATTCATAAAATGGAGCAAGGAAAT 186  
QY 1061 GTGCCCCCAGGTCTGAGGTCTGATCACCCTCAGGAAGGGGAAAGATTTCTGCAAT 1120  
Db 187 GTGCCCCCAGGTCTGAGGTCTGATCACCCTCAGGAAGGGGAAAGATTTCTGCAAT 246  
QY 1121 GAGAAATGCTGGAGTCCGTTCCACCCCTCCGAAAGATGACTCCCGTACTCCAGGTT 1180  
Db 247 GAGAAATGCTGGAGTCCGTTCCACCCCTCCGAAAGATGACTCCCGTACTCCAGGTT 306  
QY 1181 AAGAGTGGGTTCAACAGCCTGTCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 1240  
Db 307 AAGAGTGGGTTCAACAGCCTGTCATCCCAAGCCACTAAGTCCAGATTCAGAGCCTCC 366  
QY 1241 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGATTTTCAGGCACTCCGAAAGAG 1300  
Db 367 AGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGATTTTCAGGCACTCCGAAAGAG 426  
QY 1301 ACCTGCGTGAATGTCAGAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACACAG 1360  
Db 427 ACCTGCGTGAATGTCAGAGACAGTCTATCCAATGGAGCGTCTCTTGGCCAAACACAG 486  
QY 1361 GTGTTTCACATCAGTCTGCTTCCGTTGCTCTTATGCAACCAAACTCAGTCTAGGAACA 1420  
Db 487 GTGTTTCACATCAGTCTGCTTCCGTTGCTCTTATGCAACCAAACTCAGTCTAGGAACA 546  
QY 1421 TATGATCTTTTACATGCAAGATCTATTGTAAGCCTCATTCAATCAACTTTTAAATCT 1480  
Db 547 TATGATCTTTTACATGCAAGATCTATTGTAAGCCTCATTCAATCAACTTTTAAATCT 606  
QY 1481 AAGGGCAACTATGATGAAGGCTTTGGGCACAGACCACCAAGGATCTATGGGCAAGCAAA 1540  
Db 607 AAGGGCAACTATGATGAAGGCTTTGGGCACAGACCACCAAGGATCTATGGGCAAGCAAA 666  
QY 1541 AATGAAAAAAGAGATTTTGGAGAGCAGCCAGCTTGCAAAATGCAAGGGAGACCCCT 1600  
Db 667 AATGAAAAAAGAGATTTTGGAGAGCAGCCAGCTTGCAAAATGCAAGGGAGACCCCT 726  
QY 1601 CACAGCCAGGGGTAGAAGATGCCCTATTGTAAGTGGGTGCTCTGCTGCAAGTATG 1660  
Jo 727 CACAGCCAGGGGTAGAAGATGCCCTATTGTAAGTGGGTGCTCTGCTGCAAGTATG 786  
QY 1661 GAAGCCAGGCTCTCTCAGCAGGAGAGGAGACCAAGCCAGCTGAAACCAAGAAAGCTG 1720  
Db 787 GAAGCCAGGCTCTCTCAGCAGGAGAGGAGACCAAGCCAGCTGAAACCAAGAAAGCTG 846  
QY 1721 AGGATCGCTGGCCACCCCTCCTGAACTTGGAAGTTGAGGAGTCCCTTGAGGAGG 1780  
Db 847 AGGATCGCTGGCCACCCCTCCTGAACTTGGAAGTTGAGGAGTCCCTTGAGGAGG 906  
QY 1781 ATCAAAATGTCAGAAAGCCCAATGGCTCTCTGAAAGCAAAATCAGCAAGCCCAAGTTCT 1840  
Db 907 ATCAAAATGTCAGAAAGCCCAATGGCTCTCTGAAAGCAAAATCAGCAAGCCCAAGTTCT 966  
QY 1841 GAGGATGTCGATCTAGATCTGAAGAGCTTAAGACATCTTCTTCACTGAAGAAAGAGC 1900  
Db 967 GAGGATGTCGATCTAGATCTGAAGAGCTTAAGACATCTTCTTCACTGAAGAAAGAGC 1026  
QY 1901 CCCCCATTCTAGTACAGCTTCATTTCAAAGCCTCTGTCAGAGCCCAAAACTGTG 1960  
Db 1027 CCCCCATTCTAGTACAGCTTCATTTCAAAGCCTCTGTCAGAGCCCAAAACTGTG 1086  
QY 1961 TCCCCACCTATCAGAAAGGCTGAGCATGTGTCAGAGCAGAGTGAAGAGTCTGTGGTGA 2020  
Db 1087 TCCCCACCTATCAGAAAGGCTGAGCATGTGTCAGAGCAGAGTGAAGAGTCTGTGGTGA 1146  
QY 2021 AGAGTTGCAGAAAGGAAACAACTGGAAAAATGCCAAGCTTCTAAGAAAGATGGAAATGTG 2080  
Db 1147 AGAGTTGCAGAAAGGAAACAACTGGAAAAATGCCAAGCTTCTAAGAAAGATGGAAATGTG 1206  
QY 2081 GGAAAAACAACTGCCAAAAACAAAGATCTTAAGCAGAGACAGAGGAAAGAGAAATAGGAA 2140  
Db 1207 GGAAAAACAACTGCCAAAAACAAAGATCTTAAGCAGAGACAGAGGAAAGAGAAATAGGAA 1266



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 13:54:31 ; Search time 7185 Seconds  
(without alignments)  
12532.788 Million cell updates/sec

Title: US-09-890-549-16  
Perfect score: 3705  
Sequence: 1 ggcgcgggcagtaggtg.....atttgaaaaaaaaaaaaa 3705

Scoring table: Oligo NUC  
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database :

EST.\*

1: em\_estba.\*

2: em\_estnu.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_htc.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_htc.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vtt.\*

21: em\_gss\_fun.\*

22: em\_gss\_nam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	838	22.6	1075	12	BM909158	BM909158 AGENCOURT
2	746	20.1	884	14	CD171644	CD171644 AGENCOURT
3	715	19.3	728	13	BU631767	BU631767 UI-H-FLO-
4	714	19.3	777	28	AQ314676	AQ314676 RPT11-10

5	713	19.2	785	10	BG740815	BG740815 602633701
6	695	18.8	719	14	CA412726	CA412726 UI-H-EZO-
7	682	18.4	745	14	CB055210	CB055210 NISC_gm08
8	677	18.3	1142	12	BM914155	BM914155 AGENCOURT
9	676	18.2	759	14	CA447584	CA447584 UI-H-EIO-
10	676	18.2	757	14	CB988778	CB988778 AGENCOURT
11	670	18.1	695	13	BU622584	BU622584 UI-H-FLI-
12	661	17.8	711	14	CA447848	CA447848 UI-H-FLI-
13	658	17.8	787	13	BU961039	BU961039 AGENCOURT
14	647	17.5	880	12	BI871114	BI871114 UI-H-EIO-
15	645	17.4	645	12	BM786315	BM786315 AGENCOURT
16	644	17.4	647	14	CA432729	CA432729 UI-H-FLI-
17	639	17.2	989	12	BM552304	BM552304 AGENCOURT
18	638	17.2	781	10	BG676749	BG676749 602823316
19	638	17.2	918	13	BU850203	BU850203 AGENCOURT
20	637	17.2	724	12	BM980979	BM980979 UI-CF-EN1
21	635	17.1	647	13	BQ574891	BQ574891 UI-H-E21-
22	635	17.1	971	10	BE410108	BE410108 601302288
23	633	17.1	951	10	BF528886	BF528886 602043349
24	624	16.8	679	9	AL048162	AL048162 DKF2588K
25	620	16.7	725	13	BU740444	BU740444 UI-E-E01-
26	619	16.7	701	14	CD365699	CD365699 UI-H-F72-
27	618	16.7	695	14	CA420719	CA420719 UI-H-FGO-
28	613	16.5	767	12	BM981550	BM981550 UI-CF-EN1
29	611	16.5	931	12	BM449793	BM449793 AGENCOURT
30	606	16.4	635	14	CB050199	CB050199 NISC_g715
31	605	16.3	612	13	BQ632336	BQ632336 1125f12.X
32	589	15.9	643	12	BM999833	BM999833 UI-H-DP0-
33	589	15.9	643	12	BQ000468	BQ000468 UI-H-DP0-
34	586	15.8	634	10	BE729955	BE729955 601562433
35	585	15.8	759	12	BM977831	BM977831 UI-CF-EN1
36	584	15.7	643	13	BQ632628	BQ632628 1125f12.Y
37	582	15.7	724	14	CB055211	CB055211 NISC_gm08
38	577	15.6	1120	10	BG120125	BG120125 602153491
39	576	15.5	810	12	BM721370	BM721370 UI-E-E01-
40	574	15.5	574	9	AW051405	AW051405 w995b02.X
41	571	15.4	785	14	CB990409	CB990409 AGENCOURT
42	570	15.4	582	13	BQ772309	BQ772309 UI-H-E21-
43	569	15.4	582	12	BM970921	BM970921 UI-CF-EC1
44	566	15.3	837	13	BX365807	BX365807 8X365807
45	562	15.3	629	12	BQ020266	BQ020266 UI-H-ED0-

ALIGNMENTS

RESULT 1  
BM909158  
LOCUS  
AGENCOURT\_6611857 NIH\_MGC\_98 Homo sapiens cDNA clone IMAGE:5431785  
S, mRNA sequence, 1075 bp linear EST 12-MAR-2002  
BM909158  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 1075)  
NIF-MGC <http://mgc.ncl.nih.gov/>.  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLM1904 row: j column: 10  
High quality sequence stop: 696.





Db 267 ATCAGGAAGGCTGAGCATGTCAGAGCAGAGTGAAGATCTGTGGTGGAGAGTTGCA 326  
QY 2030 GAAAGGAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAGATGGGAATGTGGGAAACA 2089  
Db 327 GAAAGGAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAGATGGGAATGTGGGAAACA 386  
QY 2090 ACCTGGCAAAACAAGATCTTAAAGGAGAGACAGGGAAGAGATGAAGAGGTCATAGT 2149  
Db 387 ACCTGGCAAAACAAGATCTTAAAGGAGAGACAGGGAAGAGATGAAGAGGTCATAGT 446  
QY 2150 TTGGAGATGGAGATGAGATCTTTGAGAAATGGTGAGAAATGGTGAGAAATGGTGAG 2209  
Db 447 TTGGAGATGGAGATGAGATCTTTGAGAAATGGTGAGAAATGGTGAGAAATGGTGAG 2269  
QY 2210 AGCTTCTTCAAAACAACATCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCAGTTTGT 506  
Db 507 AGCTTCTTCAAAACAACATCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCAGTTTGT 566  
QY 2270 GACAACACCTTTGCTGAAGAAATCACTACTCAGAAATCCCAAGGATGTGGAACTC 2329  
Db 567 GACAACACCTTTGCTGAAGAAATCACTACTCAGAAATCCCAAGGATGTGGAACTC 626  
QY 2330 TGGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTAT 2389  
Db 627 TGGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTAT 686  
QY 2390 TATGATGAGATGAGATGAGATGCAAAATGCAATGATGCTGGGCCCTTAATTCATG 2449  
Db 687 TATGATGAGATGAGATGAGATGCAAAATGCAATGATGCTGGGCCCTTAATTCATG 746  
QY 2450 TTAGTGTAGGAGCCACTGCCCTTT 2475  
Db 747 TTAGTGTAGGAGCCACTGCCCTTT 772

RESULT 3  
LOCUS BU631767/c  
DEFINITION UI-H-FLO-bdp-b-01-0-UI.s1 NCI CGAP\_FLO Homo sapiens cDNA clone  
ACCESSION BU631767  
VERSION 1  
KEYWORDS 728 bp mRNA linear EST 23-SEP-2002  
SOURCE UI-H-FLO-bdp-b-01-0-UI 3', mRNA sequence.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 728)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue procurement: James Martin  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLYA=Yes

Location/Qualifiers  
1..728  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FLO-bdp-b-01-0-UI"  
/tissue\_type="Cell lines"  
/dev\_stages="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI\_CGAP\_FLO"

/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_FLO is a cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GAGGTGGTG. The cell line was provided by Dr James Martin from University of Iowa.  
TAG\_LIF=UI-H-FLO  
TAG\_TISSUE=Human Chondrosarcoma Grade 3 cell line mix  
TAG\_SEQ=GAGGTGGTG"

BASE COUNT 156 a 180 c 123 g 269 t  
ORIGIN

Query Match 19.3%; Score 715; DB 13; Length 728;  
Best Local Similarity 100.0%; Pred. No. 1.7e-164;  
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1855 AGATCTGAAGAAGCTAAGACGATCTTCTCACTGAAGGAAGAACGCCGCCATTCACTCT 1914  
Db 728 AGATCTGAAGAAGCTAAGACGATCTTCTCACTGAAGGAAGAACGCCGCCATTCACTCT 669  
QY 1915 AGCAGCTTCATTTCAAAAGCACCCTCTGTCAAGAGCCCAAAAACCTGTGTCCCACTATCAG 1974  
Db 668 AGCAGCTTCATTTCAAAAGCACCCTCTGTCAAGAGCCCAAAAACCTGTGTCCCACTATCAG 609  
QY 1975 GAAAGCTGGAGCATGTGAGAGAGAGTGAAGAGTCTGTGGTGGAGAGTGTGAGAAAG 2034  
Db 608 GAAAGCTGGAGCATGTGAGAGAGAGTGAAGAGTCTGTGGTGGAGAGTGTGAGAAAG 549  
QY 2035 GAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAAATGGGAATGTGGGAAACAACTG 2094  
Db 548 GAAACAAGTGGAAATGCCAAGGCTTCTAAGAGAAATGGGAATGTGGGAAACAACTG 489  
QY 2095 GCAAAACAAGATCTTAAAGGAGAGACAGGGAAGAGAGTGAAGAGTGTGAGTGTGGA 2154  
Db 488 GCAAAACAAGATCTTAAAGGAGAGACAGGGAAGAGAGTGAAGAGTGTGAGTGTGGA 429  
QY 2155 GATGGAGATGAGAAATCTTGTAGAAAATGGTGCAGACTCCGATCAAGATGATACAGCTT 2214  
Db 428 GATGGAGATGAGAAATCTTGTAGAAAATGGTGCAGACTCCGATCAAGATGATACAGCTT 369  
QY 2215 CCTCAACACAAATCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCGAGTTTGTAGACAA 2274  
Db 368 CCTCAACACAAATCTCCCAAGAACCCCAAGTCTCTGAATTTGGTCGAGTTTGTAGACAA 309  
QY 2275 CACCTTTGCTGAAGAATTCACACTACTCAGATCAGAAATCCAGGATGTGGAATCTGGGA 2334  
Db 308 CACCTTTGCTGAAGAATTCACACTACTCAGATCAGAAATCCAGGATGTGGAATCTGGGA 249  
QY 2335 GGGAGAAGTGGTCAAGAGAGTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTATTATGA 2394  
Db 248 GGGAGAAGTGGTCAAGAGAGTCTCTGTGGAAGAACAGATGAAGAGAAATCGGTATTATGA 189  
QY 2395 TGAGGATGAGGATGAAGAGTGAAGATGCAAAATTCATGCTGGGCCCTTAAATTCATGTAGT 2454  
Db 188 TGAGGATGAGGATGAAGAGTGAAGATGCAAAATTCATGCTGGGCCCTTAAATTCATGTAGT 129  
QY 2455 GTTAGCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATAAGCAGGTATCCAGCATCA 2514  
Db 128 GTTAGCGAGCCACTGCCCTTTGTCAAAATGTGATGCACATAAGCAGGTATCCAGCATCA 69  
QY 2515 AATGTAATTTACTTGAAGTAACTTTGGAAAGAAATTCCTTCTTAAATCAAAAA 2569  
Db 68 AATGTAATTTACTTGAAGTAACTTTGGAAAGAAATTCCTTCTTAAATCAAAAA 14

RESULT 4	777 bp	DNA	linear	GSS 04-MAY-1999
LOCUS	AQ314676			
DEFINITION	RPCI11-103F24_TV RPCI-11 Homo sapiens genomic clone RPCI-11-103F24; genomic survey sequence.			
ACCESSION	AQ314676			
VERSION	AQ314676.1	GI:4046139		
KEYWORDS	GSS.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.			
TITLE	Use of human BAC End Sequences for Sequence-Ready Map Building			
JOURNAL	Unpublished			
COMMENT	Other GSSs: RPCI11-103F24.TJ Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources ( <a href="http://bacpac.med.buffalo.edu/ordering">http://bacpac.med.buffalo.edu/ordering</a> ) or from Research Genetics (info@resgen.com). BAC end search page: <a href="http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html">http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html</a> Seq primer: T7 Class: BAC ends.			
FEATURES	Location/Qualifiers			
source	1..777			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="GBS:753911"			
	/db_xref="taxon:9606"			
	/clone="RPCI-11-103F24"			
	/sex="Male"			
	/cell_type="Lymphocytes"			
	/clone_lib="RPCI-11"			
	/notes="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"			
BASE COUNT	164 a 199 c 139 g 275 t			
ORIGIN				
Query Match	19.3%;	Score 714;	DB 28;	Length 777;
Best Local Similarity	99.9%;	Pred. No. 2.9e-164;		
Matches 764;	Conservative	0;	Mismatches 1;	Indels 0; Gaps 0;
QY	1775	GAAGGGATCAAAATGTCAAAGCCCAATGGCTCTCTGAAGCGAAATCAGCAAGCCCGAA	1834	
Db	777	GAAGGGATCAAAATGTCAAAGCCCAATGGCTCTCTGAAGCGAAATCAGCAAGCCCGAA	718	
QY	1835	GTTCTCTGAGGATGTCGATCTAGATCTGAAGAGCTTAGACCATCTTCTTCTACTGAAGGAA	1894	
Db	717	GTTCTCTGAGGATGTCGATCTAGATCTGAAGAGCTTAGACCATCTTCTTCTACTGAAGGAA	1894	
QY	1895	AGAAGCCGCCCATTCACGTGTAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAA	1954	
Db	657	AGAAGCCGCCCATTCACGTGTAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAA	598	
QY	1955	ACTGTGTCCCACTATCAGAAAGGCTGGAGCATGTGCAGCAGAGTGAAGAGTCTGTG	2014	
Db	597	ACTGTGTCCCACTATCAGAAAGGCTGGAGCATGTGCAGCAGAGTGAAGAGTCTGTG	518	
QY	2015	GGTGGAGAGTTGCAGAAAGGAAACAAGTGGAATGCCAAGGCTTCTTAAGAAGATGGG	2074	
Db	537	GGTGGAGAGTTGCAGAAAGGAAACAAGTGGAATGCCAAGGCTTCTTAAGAAGATGGG	478	



QY	3418	GTGAAATTAAACCAACTATAGCCCTTTTCTTGGGATGATTTTCTAGCTTAAGGTTTG	3477
Db	299	GTGAAATTAAACCAACTATAGCCCTTTTCTTGGGATGATTTTCTAGCTTAAGGTTTG	240
QY	3478	GGACATTTATAACCTTGAGTCATTTGTGTACACAGTTGATTTCCAAATTTGATGGATG	3537
Db	239	GGACATTTATAACCTTGAGTCATTTGTGTACACAGTTGATTTCCAAATTTGATGGATG	180
QY	3538	GGAGGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGTACTGATTTATAGAGTTTACG	3597
Db	179	GGAGGGAGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGTACTGATTTATAGAGTTTACG	120
QY	3598	TTTAATATTTTTTATAGATGTAAACACATTTCTGCTTTCTTAACTTACCTAGTCTGAACA	3657
Db	119	TTTAATATTTTTTATAGAGATGTAAACACATTTCTGCTTTCTTAACTTACCTAGTCTGAACA	60
QY	3658	TTTTTAATCAATAAGATTTTAAATTAATAATTGAA	3692
Db	59	TTTTTAATCAATAAGATTTTAAATTAATAATTGAA	25
RESULT 7			
CB055210/c			
LOCUS	CB055210	745 bp	mRNA linear EST 17-JAN-2003
DEFINITION	NISC_gm08e09.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone		
ACCESSION	IMAGE:3291761 3', mRNA sequence.		
VERSION	CB055210		
KEYWORDS	EST		
SOURCE	CB055210.1 GI:27793497		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 745)		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
COMMENT	Unpublished Contact: Robert Strausberg, Ph.D. Email: CGAPbs-w@mail.nih.gov cDNA Library Preparation: CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: infoimage.llnl.gov Plate: LLN8061 Row: J Column: 18 Seq primer: -21M13 forward primer (ABI). Location/Qualifiers 1..745 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:3291761" /tissue_type="Glioblastoma (pooled)" /lab_host="DH10B" /clone_lib="NCI CGAP Brn23" /note="Organ: brain; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCGAAGTGGAGCGGCATATCTTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."		
BASE COUNT	261 a	126 c	130 g 228 t
ORIGIN			
Query Match	18.4%	Score 682;	DB 14; Length 745;
Best Local Similarity	99.9%	Pred. No. 1.8e-156;	

Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LCM1999 row: k column: 09  
 High quality sequence stop: 678.

## FEATURES

Location/Qualifiers  
 1. .1142  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5479040"  
 /tissue\_type="astrocytoma grade IV, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 98"  
 /note="Organ: brain; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH MGC Library."  
 376 a 282 c 276 g 208 t

BASE COUNT  
 ORIGIN

Query Match  
 Best Local Similarity 18.3%; Score 677; DB 12; Length 1142;  
 Matches 727; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

347 GTGTTAAAGAGAAGTGGGAGAACCCAGGGCTGGGACGAGTCTCACAGACTCTCTA 406  
 Db 1 GTGTTAAGAGAAGTGGGAGAACCCAGGGCTGGGACGAGTCTCACAGACTCTCTA 60

407 CGGAACAGCAGCCTGAGATTAGGCACAGACAGACCATCTCTCTGCTGAAGTGACAAGC 456  
 Db 61 CGGAACAGCAGCCTGAGATTAGGCACAGACAGACCATCTCTCTGCTGAAGTGACAAGC 120

467 CACGCTGCTTCTGAGGCCAAGCTGACAAAGAACAAATCCACCCAGACTTAGACTC 536  
 Db 121 CACGCTGCTTCTGAGGCCAAGCTGACAAAGAACAAATCCACCCAGACTTAGACTC 180

527 AGGTCACCTCTGAGCCCTCGTTACGGTTCGATATCCCCACATCAAGGACGTTGAGGAT 586  
 Db 181 AGGTCACCTCTGAGCCCTCGTTACGGTTCGATATCCCCACATCAAGGACGTTGAGGAT 240

587 CTTAAAGACCACTCAACAGAAAAGTAAATAATGGAAATTTGCTAGGAGATCCAGGCAT 646  
 Db 241 CTTAAAGACCACTCAACAGAAAAGTAAATAATGGAAATTTGCTAGGAGATCCAGGCAT 300

647 GAAGTAGAAAATCAGAAATCAGTGAATAACACAGATGCTTCGGGCAAAATAGAGAAATAT 706  
 Db 301 GAAGTAGAAAATCAGAAATCAGTGAATAACACAGATGCTTCGGGCAAAATAGAGAAATAT 360

707 AATGTTCCGCTGACAGGCTTAGATGATGTTTGAAGAGGTGAACCAACTCAAACTAAG 766  
 Db 361 AATGTTCCGCTGACAGGCTTAGATGATGTTTGAAGAGGTGAACCAACTCAAACTAAG 420

767 ATTCTCGGGCCCAAGCGCAAGTGAAGTGAAGCAAGATCTCTGAAACAGCTATTCT 826  
 Db 421 ATTCTCGGGCCCAAGCGCAAGTGAAGTGAAGCAAGATCTCTGAAACAGCTATTCT 480

827 CTAGATGACTGGAATAGGCCAGGTGAGTGTGTCATCTTACATTTGACTCGGAGAA 886  
 Db 481 CTAGATGACTGGAATAGGCCAGGTGAGTGTGTCATCTTACATTTGACTCGGAGAA 540

887 AATGAGAGTAGCAAAATCTGAACTTCCACCGCTCTCAGAAACCTCTATAAGGATCGA 946  
 Db 541 AATGAGAGTAGCAAAATCTGAACTTCCACCGCTCTCAGAAACCTCTATAAGGATCGA 600

QY 947 ATGGCCAGTACCAGGAGCTGTGTCACAAACAAAGCAGCTCAACCAACTATACAAATGAG 1006  
 Db 501 ATGGCCAGTACCAGGAGCTGTGTCACAAACAAAGCAGCTCAACCAACTATACAAATGAG 660

QY 1007 CTGAAAGCCAGTGTGTCGGAATCAAAATTCATAAAATGAGCAAAAGGAGAAATGTGCC 1066  
 Db 661 CTGAAAGCCAGTGTGTCGGAATCAAAATTCATAAAATGAGCAAAAGGAGAAATGTGCC 720

QY 1067 CCAGGTCC 1074  
 Db 721 CCAGGTCC 728

RESULT 9  
 CA447584 759 bp mRNA linear EST 08-NOV-2002  
 LOCUS  
 DEFINITION  
 UI-H-E10-aye-b-12-0-UI.s1 NCI CGAP\_E10 Homo sapiens cDNA clone  
 UI-H-E10-aye-b-12-0-UI 3', mRNA sequence.  
 CA447584  
 VERSION  
 CA447584.1 GI:24812004  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 759)  
 AUTHORS  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>  
 TITLE  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL  
 Unpublished  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabps@mail.nih.gov](mailto:cgabps@mail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@iowa.edu](mailto:bento-soares@iowa.edu)  
 The following repetitive elements were found in this cDNA  
 sequence: 1-52, SAT rich#Low\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

## FEATURES

Location/Qualifiers  
 1. .759  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-E10-aye-b-12-0-UI"  
 /tissue\_type="Chondrosarcoma"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP\_E10"  
 /note="Organ: Left Pelvis; Vector: p7T3-pac (Pharmacia)  
 with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP E10 is a cDNA library containing the following  
 tissue(s): Chondrosarcoma. The library was constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,  
 6:791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into p7T3-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tag for this library is ACATTGCAC.  
 TAG LIB=UI-H-E10  
 TAG TISSUE=chondrosarcoma  
 TAG\_SEQ=ACATTGCAC"  
 BASE COUNT 266 a 129 c 132 g 229 t 3 others  
 ORIGIN

Query Match  
 Best Local Similarity 18.2%; Score 676; DB 14; Length 759;  
 Matches 99.9%; Pred. No. 5e-155;

Matches	726;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	2979	ATGCTGCTCTCTAACTACATCTCTGAACCTGACCTCTGAGGTATATATAACAGAGC	3038						
Db	730	ATGCTGCTCTCTAACTACATCTCTGAACCTGACCTCTGAGGTATATATAACAGAGC	671						
QY	3039	ACTTTTGTAGGCAATTTGAAAAACCAACCTACACTCTTCGGTCTTAGAGAGATCTGCTGT	3098						
Db	670	ACTTTNTAGGCAATTTGAAAAACCAACCTACACTCTTCGGTCTTAGAGAGATCTGCTGT	611						
QY	3099	CTCCCAAAATAGCTTTGTATCTGCGAGTGAATTTACTGTACTCTCAAAATGATTGCTTCT	3158						
Db	610	CTCCCAAAATAGCTTTGTATCTGCGAGTGAATTTACTGTACTCTCAAAATGATTGCTTCT	551						
QY	3159	TTTCTGCTGATCTGCTCTCTCATAATTAAGCTGAAGCTGAATTTTAGTATATCT	3218						
Db	550	TTTCTGCTGATCTGCTCTCTCATAATTAAGCTGAAGCTGAATTTTAGTATATCT	491						
QY	3219	TGGGATCACTGCTCCCTCATCTTCCTGTTTAGAGCAAGTGAAGTGAAGGGAAG	3278						
Db	490	TGGGATCACTGCTCCCTCATCTTCCTGTTTAGAGCAAGTGAAGTGAAGGGAAG	431						
QY	3279	AGAAGAACTGCTTTACACCACTTGAGCTCAGACTCTAAACCTGATTTCCCTTATG	3338						
Db	430	AGAAGAACTGCTTTACACCACTTGAGCTCAGACTCTAAACCTGATTTCCCTTATG	371						
QY	3339	ATGCTCCCTTTTGAGACACTTAATTTTAAATTAATTAATTAATTAATTAATTAAT	3398						
Db	370	ATGCTCCCTTTTGAGACACTTAATTTTAAATTAATTAATTAATTAATTAATTAAT	311						
QY	3399	TTATCACAGTATCTCAGGCTGAATTTAAACCAACTATAGGCTTTTCTTGGGATGAT	3458						
Db	310	TTATCACAGTATCTCAGGCTGAATTTAAACCAACTATAGGCTTTTCTTGGGATGAT	251						
QY	3459	TTCTAGTCTTAAGGTTTGGGACATTATAAATTTGAGTACATTTTGTGTACACAGTTGAT	3518						
Db	250	TTCTAGTCTTAAGGTTTGGGACATTATAAATTTGAGTACATTTTGTGTACACAGTTGAT	191						
QY	3519	ATTCAAATTTGATGGTGGGAGGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGACT	3578						
Db	190	ATTCAAATTTGATGGTGGGAGGAGGTGCTTAAAGCTGTAGGCTTTTCTTTGACT	131						
QY	3579	GCATTTATAGAGATTAGCTTTAATTTTATTTTAGAGATGTAACATTTCTCTTTAG	3638						
Db	130	GCATTTATAGAGATTAGCTTTAATTTTATTTTAGAGATGTAACATTTCTCTTTAG	71						
QY	3639	TCTTACCTAGTCTGAACATTTTATCAATAAGATTTTAAATTAATTAATTAATTAAT	3698						
Db	70	TCTTACCTAGTCTGAACATTTTATCAATAAGATTTTAAATTAATTAATTAATTAAT	11						
QY	3699	AAAAAAA 3705							
Db	10	AAAAAAA 4							

RESULT 10  
CB988778  
LOCUS  
DEFINITION  
AGENCY 13927895 NIH MGC 147 Homo sapiens cDNA clone  
IMAGE:30341684 5', mRNA sequence.  
CB988778  
VERSION  
KEYWORDS  
SOURCE  
EST.  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 767)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-re@mail.nih.gov

Tissue Procurement: Dr. Stefan Hansson  
cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help  
and advice from Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM373 row: i column: 21  
High quality sequence stop: 649.  
Location/Qualifiers  
1..767  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30341684"  
/tissue\_type="Human Placenta"  
/lab\_host="DH10B Tona"  
/clone\_lib="NIH MGC 147"  
/note="Organ: placenta; Vector: pBluescriptR; Site: 1:  
all-XhoI; Site: 2: BamH; Oligo-dT primed using primer  
5'-TTTCTTTTCTTTTCTTTT-3', size-selected for average  
insert size 2.3 kb and normalized to ROT 5. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein, in  
(NIH/NHGRI, National Institutes of Health). Note: This is  
a NIH MGC library."

FEATURES	source
BASE COUNT	256 a 168 c 195 g 148 t
ORIGIN	
Query Match	18.3%; Score 676; DB 14; Length 767;
Best Local Similarity	100.0%; Pred. No. 5e-155;
Matches	676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	16	AGGTGTTAGCAGCTTGGTCCGACAGGTGCGCTAGGTAGAGCCCGGACCTGTGACAGG	75
Db	28	AGGTGTTAGCAGCTTGGTCCGACAGGTGCGCTAGGTAGAGCCCGGACCTGTGACAGG	87
QY	76	GCTGTTAGCAGCCAGAGAAAGCGGCTTTTAGCCAGGTATTTCACTGCTGTAGACAA	135
Db	88	GCTGTTAGCAGCCAGAGAAAGCGGCTTTTAGCCAGGTATTTCACTGCTGTAGACAA	147
QY	136	GATGGATCATCTCCATTTAATAGCGGCAATGGACCTCACTATCATTAGGGTAACAG	195
Db	148	GATGGATCATCTCCATTTAATAGCGGCAATGGACCTCACTATCATTAGGGTAACAG	207
QY	196	CAAGAACTTTCTCTTGTCAACAAGAAACAAATGAGAGAGAGAGAGAGAGAGAGAG	255
Db	208	CAAGAACTTTCTCTTGTCAACAAGAAACAAATGAGAGAGAGAGAGAGAGAGAGAG	267
QY	256	GTACCAAG	315
Db	268	GTACCAAG	327
QY	316	CTCCAG	375
Db	328	CTCCAG	387
QY	376	GCTGGAGCAGAGAGTCTCACACAGAGTCTCTACCGGAAACAGCAGCAGTATAGG	435
Db	388	GCTGGAGCAGAGTCTCTCACACAGAGTCTCTACCGGAAACAGCAGCAGTATAGG	447
QY	436	AGCAGACCATCTCTCTGTTGAAGTGAAGGCAAGCAGCAGTCTCTCTGAGGCAAG	495
Db	448	AGCAGACCATCTCTCTGTTGAAGTGAAGGCAAGCAGCAGTCTCTCTGAGGCAAG	507
QY	496	AGAAGAACTTCCACCCAGATCTTAGAGTCACTGAGTCACTCTCTGAGGCTCTG	555
Db	508	AGAAGAACTTCCACCCAGATCTTAGAGTCACTGAGTCACTCTCTGAGGCTCTG	567
QY	556	TCGATATCCCCACATCAAGGACCGGTGAGGATCTTAAAGACCACTCAACAGAAAG	615

```

Db      568  TCGATATCCCATCATCAAGGAGCGGTGAGATCTTTAAAGACCACCTCAACAGAAAGTAAAAA 627
QY      616  AATGAAAAATTTCTAGGAGAAATCCAGGCATGAGTACAGAAATCAGAAATCAGTAAAA 675
Db      628  AATGAAAAATTTCTAGGAGAAATCCAGGCATGAGTACAGAAATCAGTAAAA 687
QY      676  CACAGATGCTTCGGGC 691
Db      688  CACAGATGCTTCGGGC 703

RESULT 11
BU622584/c
LOCUS      695 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-FL1-bgc-o-10-0-UI s1 NCI CGAP_Fl1 Homo sapiens CDNA clone
ACCESSION UI-H-FL1-bgc-o-10-0-UI 3', mRNA sequence.
VERSION    BU622584.1
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 695)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Cl. Bento Soares, University of Iowa
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seg primer: M13 FORWARD
POLYA=Yes

FEATURES             Location/Qualifiers
     source            1..695
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="UI-H-FL1-bgc-o-10-0-UI"
                     /tissue_type="Cell lines"
                     /dev_stage="Adult"
                     /lab_host="DH10B (Life Technologies)"
                     /clone_lib="NCI CGAP FL1"
                     /notes="Organ: Chondrosarcoma; Vector: pT7T3-Pac (Pharmacia)
                     with a modified polylinker; Site 1: EcoR I; Site 2: Not
                     I; NCI CGAP_Fl1 is a normalized cDNA library derived from
                     a pool of mRNA obtained from 4 cell lines from grade III
                     chondrosarcoma tissues. The library was constructed
                     according to Bonaldo, Lennon and Soares, Genome Research,
                     6:791-806, 1996. First strand cDNA synthesis was primed
                     with an oligo-dT primer containing a Not I site. Double
                     stranded cDNA was ligated to an EcoR I adaptor, digested
                     with Not I, and cloned directionally into pT7T3-Pac
                     vector. The oligonucleotide used to prime the synthesis of
                     first-strand cDNA contains a library tag sequence that is
                     located between the Not I site and the (dT)18 tail. The
                     sequence tag for this library is GAGTCGGTG. The cell
                     lines were provided by Dr. James Martin from the
                     University of Iowa.
                     TAG_LIB=UI-H-FL1
                     TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
                     TAG_SEQ=CAGTCGGTG"

BASE COUNT  148 a  173 c  116 g  258 t
ORIGIN

```

Query Match 18.1%; Score 670; DB 13; Length 695;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-153;  
 Matches 670; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1900  CCGCCCATTCACCTGTAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAAACCTGT 1959
Db      683  CCGCCCATTCACCTGTAGCAGCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAAACCTGT 624
QY      1960  GTCCCCACCTATCAGGAAAGGCTGGAGCATGTGTCAGAGCAGAGTGAAGAGTCTCTGGGTGG 2019
Db      623  GTCCCCACCTATCAGGAAAGGCTGGAGCATGTGTCAGAGCAGAGTGAAGAGTCTCTGGGTGG 564
QY      2020  AAGAGTTTCAGAAAGGAAACAAGTGCAGAAATGCCAAGGGTCTCTAAAGAAATGGGAATGT 2079
Db      563  AAGAGTTTCAGAAAGGAAACAAGTGCAGAAATGCCAAGGGTCTCTAAAGAAATGGGAATGT 504
QY      2080  GGGAAAAACAACCTCGCAAAAACAAGAAATCTAAAGGAGAGACAGGGAAGAGTAAGGA 2139
Db      503  GGGAAAAACAACCTCGCAAAAACAAGAAATCTAAAGGAGAGACAGGGAAGAGTAAGGA 444
QY      2140  AGGTCATAGTTTGGAGATGAGAAATCTTGTAGAAAATGGTGCAGACTCCGATGA 2199
Db      443  AGGTCATAGTTTGGAGATGAGAAATCTTGTAGAAAATGGTGCAGACTCCGATGA 384
QY      2200  AGATGATACACCTTCCTCAAAACAACATCTCCACAAGAACCCAAAGTCTCTGAATTGGTC 2259
Db      383  AGATGATACACCTTCCTCAAAACAACATCTCCACAAGAACCCAAAGTCTCTGAATTGGTC 324
QY      2260  GAGTTTGTAGACAACACCTTTGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCCAGCA 2319
Db      323  GAGTTTGTAGACAACACCTTTGCTGAAGAAATTCACCTACTCAGAAATCAGAAATCCCAGCA 264
QY      2320  TGTGGAATCTCTGGAGGAGAAAGTGTCAAAAGAGCTCTCTGTGGAGAACACAGATAAAGAG 2379
Db      263  TGTGGAATCTCTGGAGGAGAAAGTGTCAAAAGAGCTCTCTGTGGAGAACACAGATAAAGAG 204
QY      2380  AAATCGGTATTATGATGAGGATGAGATGAAGAGTGACAAATTCGAATGATGCTGGGCT 2439
Db      203  AAATCGGTATTATGATGAGGATGAGATGAAGAGTGACAAATTCGAATGATGCTGGGCT 144
QY      2440  TAAATTCATGTTAGTGTAGCGAGGACCTGCCCTTTGCAAAATGTGATGCACATAAGCA 2499
Db      143  TAAATTCATGTTAGTGTAGCGAGGACCTGCCCTTTGCAAAATGTGATGCACATAAGCA 84
QY      2500  GGTATCCCGAGCATGAATGTAATTTACTTGGAGTAACCTTTGGAAAGAAATTCCTTCTTA 2559
Db      83  GGTATCCCGAGCATGAATGTAATTTACTTGGAGTAACCTTTGGAAAGAAATTCCTTCTTA 24
QY      2560  AAATCAAAAA 2569
Db      23  AAATCAAAAA 14

CA447848 711 bp mRNA linear EST 08-NOV-2002
UI-H-B10-ayf-j-13-0-UI-s1 NCI CGAP_E10 Homo sapiens CDNA clone
UI-H-B10-ayf-j-13-0-UI 3', mRNA sequence.

CA447848
CA447848.1 GI:24812268
EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 711)
NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

```

RESULT 12  
 CA447848/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT



Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu  
The following repetitive elements were found in this cDNA sequence: 1-52, SAT rich#Low complexity (matched complement)  
Seq primer: M13 FORWARD  
POLVA=Yes.

FEATURES  
Source  
Location/Qualifiers  
1. .711  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-E10-avf-j-13-0-UI"  
/tissue\_type="Chondrosarcoma"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP E10"  
/note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP E10 is a cDNA library containing the following tissue(s): Chondrosarcoma. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACACTTGAC.  
TAG LIB=UI-H-E10  
TAG\_TISSUE=chondrosarcoma  
TAG\_SEQ=ACACTTGAC"  
BASE COUNT 248 a 123 c 123 g 216 t 1 others  
ORIGIN

Query Match 17.8%; Score 661; DB 14; Length 711;  
Best Local Similarity 100.0%; Pred. No. 2.3e-151;  
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3045 TGAGCAATTGAAAACCAACCTACACTCTTCGGTGCTTAGAGAGATCGTGCTCCCA 3104  
DB 664 TGAGCAATTGAAAACCAACCTACACTCTTCGGTGCTTAGAGAGATCGTGCTCCCA 605  
QY 3105 AATAAGCTTTGTATCTCCAGTGAATTTACTGACTCCAAATGATTCCTTTCTG 3164  
DB 604 AATAAGCTTTGTATCTCCAGTGAATTTACTGACTCCAAATGATTCCTTTCTG 545  
QY 3165 GTGATATCTGCTTCTCATTAATTACTGAAAGCTGCAATATTTAGTAATACCTTCGGGA 3224  
DB 544 GTGATATCTGCTTCTCATTAATTACTGAAAGCTGCAATATTTAGTAATACCTTCGGGA 495  
QY 3225 TCACCTGTCCCCCATCTCCCGTGTAGAGCAAGTGAAGAGTTTAAAGAGGAGAGAGAAA 3284  
DB 484 TCACCTGTCCCCCATCTCCCGTGTAGAGCAAGTGAAGAGTTTAAAGAGGAGAGAGAAA 425  
QY 3285 GAACGTCTTACACCACTTGAGCTCAGACCTCTTAAACCCCTGATTTCCCTTTGATGTC 3344  
DB 424 GAACGTCTTACACCACTTGAGCTCAGACCTCTTAAACCCCTGATTTCCCTTTGATGTC 365  
QY 3345 CCTTTTGTAGACACTAATTTTAAATCTACTAGCTCTGAAATATATGATTTTATCA 3404  
DB 364 CCTTTTGTAGACACTAATTTTAAATCTACTAGCTCTGAAATATATGATTTTATCA 305  
QY 3405 CAGTATCTCAGGTGAAATTTAAACCAACTATAGGCTTTTCTGGGATGATTTCTAG 3464  
DB 304 CAGTATCTCAGGTGAAATTTAAACCAACTATAGGCTTTTCTGGGATGATTTCTAG 245  
QY 3465 TCTTAAGGTTTGGGACATTATAAACAATTGAGTACATTTGTTACACAGTTGATTTCCA 3524  
DB 244 TCTTAAGGTTTGGGACATTATAAACAATTGAGTACATTTGTTACACAGTTGATTTCCA 185  
QY 3525 AATTGTATGATGGAGGGAGAGTGCTTAAAGCTAGGCTTTTCTTGTACTGCATT 3584

Db 184 AATTGTATGATGGAGGGAGAGGTGCTTAAAGCTAGGCTTTTCTTGTACTGCATT 125  
QY 3585 ATAGAGATTTAGCTTTAATATTTTAGAGATGTAAACACATTCTGCTTTCTAGTCTTAC 3644  
Db 124 ATAGAGATTTAGCTTTAATATTTTAGAGATGTAAACACATTCTGCTTTCTAGTCTTAC 65  
QY 3645 CTAGCTGAAACATTTTATTATTAATTAAGATTTTAAATTAATTTGAAAAA 3704  
Db 64 CTAGCTGAAACATTTTATTATTAATTAAGATTTTAAATTAATTTGAAAAA 5  
QY 3705 A 3705  
Db 4 A 4  
RESULT 13  
BU861039  
LOCUS  
DEFINITION BU861039 787 bp mRNA linear EST 16-OCT-2002  
IMAGE:6652372 5', mRNA\_sequence.  
ACCESSION BU861039  
VERSION BU861039.1 GI:24046031  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 787)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-f@mail.nih.gov  
Tissue Procurement: NCI  
CDNA Library Preparation: Michael Brownstein Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LCM2900 row: 1 column: 04  
High quality sequence stop: 617.  
Location/Qualifiers  
1. .787  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6652372"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (Tl-phage-resistant)"  
/clone\_lib="NIH MGC 126"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.2%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
5'-AAGCATGTGTATCAACGAGATGGCCATTACGGCCGG-3' and 5'-ATTCTAGAGCGGAGCGGCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-1 kb size fraction (other fractions present in NIH\_MGC 127 and NIH\_MGC 128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC Library."

FEATURES  
Source  
Location/Qualifiers  
1. .787  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6652372"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (Tl-phage-resistant)"  
/clone\_lib="NIH MGC 126"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.2%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
5'-AAGCATGTGTATCAACGAGATGGCCATTACGGCCGG-3' and 5'-ATTCTAGAGCGGAGCGGCGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-1 kb size fraction (other fractions present in NIH\_MGC 127 and NIH\_MGC 128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC Library."

BASE COUNT 269 a 152 c 217 g 148 t 1 others  
ORIGIN

Query Match 17.8%; Score 658; DB 13; Length 787;  
Best Local Similarity 100.0%; Pred. No. 1.2e-150;

Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1630 TGCTAAGGTGGTGTCTGGCTGCAAGTATGGAAGCCAGGCTCTCTCTAGCAGAGAA 1689  
Db |||||  
3 TGCTAAGGTGGTGTCTGGCTGCAAGTATGGAAGCCAGGCTCTCTCTAGCAGAGAA 62

QY 1690 GGAAGACAAGCCAGCTGAACCAAGAGAGCTGAGGATGCGCTGGCCACCCCTGAACT 1749  
Db |||||  
63 GGAAGACAAGCCAGCTGAAACCAAGAGAGCTGAGGATGCGCTGGCCACCCCTGAACT 122

QY 1750 TGGAACTTCAGGAAGTGCCTTGGAGGAGGATCAAAATGTCAAAGCCCAATGGCCTCC 1809  
Db |||||  
123 TGGAACTTCAGGAAGTGCCTTGGAGGAGGATCAAAATGTCAAAGCCCAATGGCCTCC 182

QY 1810 TGAGAGCAAAATCAGCAAGCCGAGTCTCTGAGGATGTGATCTAGATCTGAAGAAGCT 1869  
Db |||||  
183 TGAAGACAAATCAGCAAGCCGAGTCTCTGAGGATGTGATCTAGATCTGAAGAAGCT 242

QY 1870 AAGACGATCTCTTCACTGAAGGAGAAAGCCGCCATTCACCTGTAGCAGCTTCATTTCA 1929  
Db |||||  
243 AAGACGATCTCTTCACTGAAGGAGAAAGCCGCCATTCACCTGTAGCAGCTTCATTTCA 302

QY 1930 AAGCACCTCTGTCAAGAGCCCAAAACTGTCTCCCACTATCAGGAAGGCTGGAGCAT 1989  
Db |||||  
303 AAGCACCTCTGTCAAGAGCCCAAAACTGTCTCCCACTATCAGGAAGGCTGGAGCAT 362

QY 1990 GTCAAGCAGAGTGAAGAGTCTGTGGTGGAGAGTTGCGAAGGAGAAACAACTGGAA 2049  
Db |||||  
363 GTCAAGCAGAGTGAAGAGTCTGTGGTGGAGAGTTGCGAAGGAGAAACAACTGGAA 422

QY 2050 TGCCAAGGCTTCTAAGAGAAATGGGAATGTGGGAAAGCAACCTGGCAAAACAAAGATC 2109  
Db |||||  
423 TGCCAAGGCTTCTAAGAGAAATGGGAATGTGGGAAAGCAACCTGGCAAAACAAAGATC 482

QY 2110 TAAAGAGAGACAGGAGAGAGTGAAGAGTGTAGTGTGGAGATGGAGATGAGAA 2169  
Db |||||  
483 TAAAGAGAGACAGGAGAGAGTGAAGAGTGTAGTGTGGAGATGGAGATGAGAA 542

QY 2170 TCTTGTAGAAAATGGTGCAGATCCGATGAAGATGATAACAGCTTCTCTCAAACAACATC 2229  
Db |||||  
543 TCTTGTAGAAAATGGTGCAGATCCGATGAAGATGATAACAGCTTCTCTCAAACAACATC 602

QY 2230 TCCACAAAGACCAAGCTCTGAAATGGTGGAGATTTGTAGCAACACCTTTGCTGAA 2287  
Db |||||  
603 TCCACAAAGACCAAGCTCTGAAATGGTGGAGATTTGTAGCAACACCTTTGCTGAA 660

RESULT 14  
BI871114  
LOCUS  
DEFINITION  
60394291F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5404054 5',  
mRNA sequence.  
BI871114  
BI871114.1 GI:16044787  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 880)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM12030 row: f column: 23

High quality sequence stop: 799.  
Location/Qualifiers  
1. .880  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5404054"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH108 (phage-resistant)"  
/clone\_lib="NIH\_MGC\_90"  
/note="Organ: liver; Vector: pCMV-Sport6; Site: 1; Not1;  
Site: 2; Salt; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

BASE COUNT 308 a 136 c 181 g 255 t  
ORIGIN

Query Match 17.5%; Score 647; DB 12; Length 880;  
Best Local Similarity 99.9%; Pred No. 5,1e-148;  
Matches 697; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2124 GGAAGAGAAGTGAAGAGTCTAGCTTTGGAGATGGAGATGAGAAATCTGTAGAAAAATG 2183  
Db |||||  
1 GGAAGAGAAGTGAAGAGTCTAGCTTTGGAGATGGAGATGAGAAATCTGTAGAAAAATG 60

QY 2184 GTGCAGATCCGATGAAGATGATACAGCTTCTCTCAAACAACAAATCTCCACAAGAACCCA 2243  
Db |||||  
61 GTGCAGATCCGATGAAGATGATACAGCTTCTCTCAAACAACAAATCTCCACAAGAACCCA 120

QY 2244 AGTCTCTGAATTTGGTCGAGTTTGTAGACAACACCTTTGCTGGAAGATTCACCTACTCAGA 2303  
Db |||||  
121 AGTCTCTGAATTTGGTCGAGTTTGTAGACAACACCTTTGCTGGAAGATTCACCTACTCAGA 180

QY 2304 ATCAGAAATCCAGGATGTGGAACCTCTGGGAGGAGAAAGTGTCAAAGAGCTCTCTGTGG 2363  
Db |||||  
181 ATCAGAAATCCAGGATGTGGAACCTCTGGGAGGAGAAAGTGTCAAAGAGCTCTCTGTGG 240

QY 2364 AAGAACAGATGAAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGACAAATG 2423  
Db |||||  
241 AAGAACAGATGAAGAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGACAAATG 300

QY 2424 CAATGATGCTGGGCTTAAATTCATGTTAGTGTAGGAGCCACTGCCCTTTGTCAAAT 2483  
Db |||||  
301 CAATGATGCTGGGCTTAAATTCATGTTAGTGTAGGAGCCACTGCCCTTTGTCAAAT 360

QY 2484 GTGATGCACATGAAGCAGTATCCAGCATGAAATGTAATTTACTTGGAGTAACCTTTGGA 2543  
Db |||||  
361 GTGATGCACATGAAGCAGTATCCAGCATGAAATGTAATTTACTTGGAGTAACCTTTGGA 420

QY 2544 AAAGAATTCCTTCTTAAATCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACATTTAAAT 2603  
Db |||||  
421 AAAGAATTCCTTCTTAAATCAAAAACAAAACAAAACAAAACAAAACAAAACAAAACATTTAAAT 480

QY 2604 ACTAGAGATTAACCTTAAATTCATTTTAGCAGTATGATATGCAATAGTGTCTGT 2663  
Db |||||  
481 ACTAGAGATTAACCTTAAATTCATTTTAGCAGTATGATATGCAATAGTGTCTGT 540

QY 2664 AAGCTTTGTAAGTGGGCAATATCCACCTGATTAAGCCAGATTTCTACTGTTATCCCA 2723  
Db |||||  
541 AAGCTTTGTAAGTGGGCAATATTTCCACCTGATTAAGCCAGATTTCTACTGTTATCCCA 600

QY 2724 AAAGGCAATATTAAAGGTAGATGATGATATGATATGTTTACACACTATTTTTGAAT 2783  
Db |||||  
601 AAAGGCAATATTAAAGGTAGATGATGATATGATATGTTTACACACTATTTTTGAAT 660

QY 2784 TAGAGAACATACAGAAGGAATTTAGGGGCTTAAACATT 2821  
Db |||||  
661 TAGAGAACATACAGAAGGAATTTAGGGGCTTAAACATT 698

RESULT 15  
BM786315

LOCUS BM786315 645 bp mRNA linear EST 05-MAR-2002  
DEFINITION K-ST0065015 S9SNU601 Homo sapiens cDNA clone S9SNU601-36-B02 5',  
mRNA sequence.  
ACCESSION BM786315  
VERSION BM786315.1 GI:19134547  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 645)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Ch,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 36 row: B column: 02  
High quality sequence stop: 645.  
FEATURES  
source  
1. 645  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S9SNU601-36-B02"  
/sex="M"  
/tissue\_type="Ascites"  
/cell\_type="Epithelial"  
/lab\_host="SNU-601"  
/lab\_host="Top10P"  
/clone\_lib="S9SNU601"  
Note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI;  
Site 2: XhoI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with Cabacoo acid pyrophosphatase (TAP). The decapped  
inact mRNA was ligated with DNA-RNA linker including Sfil  
site by treatment of T4 RNA ligase and the first strand  
cDNA was synthesized with Superscript II using Sfil  
oligo-dT primer. After first strand synthesis, RNA was  
degraded by NaOH treatment and cDNA was amplified by PCR  
reaction. The PCR products were digested with Sfil and  
cloned into DrallI- digested pME18S-FL3 vector. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10P, by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."  
BASE COUNT 198 a 110 c 116 g 221 t  
ORIGIN  
Query Match 17.4%; Score 645; DB 12; Length 645;  
Best Local Similarity 100.0%; Pred. No. 1.9e-147;  
Matches 645; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2658 TGCTGTAAGCGTTGTAACTGGGGAAATATTCACCTGATATAGCCAGATTCTACTGTA 2717  
DB 1 TGCTGTAAGCGTTGTAACTGGGGAAATATTCACCTGATATAGCCAGATTCTACTGTA 60  
QY 2718 TTCCCAAAAGGCAATATTAAAGGTAGATAGATGATTAGTAGTATATTGTACACATTTT 2777  
DB 61 TTCCCAAAAGGCAATATTAAAGGTAGATAGATGATTAGTAGTATATTGTACACATTTT 120  
QY 2778 TGGAAATTAGAGAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAATGCATT 2837  
DB 121 TGGAAATTAGAGAACATACAGAGGAATTTAGGGGCTTAAACATTACGACTGAATGCATT 180  
QY 2838 TAGTATTAAGGGCACAGTTTGTTATATTTTAAATCAATCAATTTAATTTTTTAGTATT 2897

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 15:54:36 ; Search time 199 seconds  
(without alignments)  
8217.711 Million cell updates/sec

Title: US-09-890-549-16  
Perfect score: 3705  
Sequence: 1 ggcgcaggagcagtaggtg.....atttgaaaaa..... 3705  
Scoring table: OLIGO NUC  
Gapop 60.0, Gapext 60.0  
Searched: 569978 seqs, 220691566 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2565	69.2	2905	4	US-09-620-312D-45
2	362	9.8	586	3	US-09-385-982-139
3	276	7.4	279	4	US-09-016-434-493
4	248	6.7	413	4	US-09-016-434-680
5	119	3.2	238	4	US-09-016-434-181
6	24	0.6	2022	2	US-08-937-540-7
7	24	0.6	3138	1	US-07-867-106-4
8	23	0.6	199	1	US-08-330-108-4
9	23	0.6	199	5	PCT-US92-10087-4
10	23	0.6	2047	3	US-09-345-468-1
11	23	0.6	2047	4	US-09-414-453A-1
12	23	0.6	2047	4	US-09-832-312-1
13	23	0.6	3556	3	US-09-276-531-8
14	22	0.6	162450	4	US-09-345-882-1
15	21	0.6	720	2	US-08-064-694-1
16	21	0.6	966	2	US-08-766-738-2
17	21	0.6	966	4	US-09-016-610-2
18	21	0.6	985	4	US-09-016-434-1223
19	21	0.6	1444	4	US-09-371-671B-1
20	21	0.6	1975	1	US-08-484-105-11
21	21	0.6	1975	1	US-08-484-105-11
22	21	0.6	2020	4	US-09-347-650-3
23	21	0.6	3157	6	5198347-3
24	21	0.6	246240	2	US-08-724-394A-20
25	21	0.6	246240	2	US-08-724-394A-21
26	21	0.6	246240	2	US-08-724-394A-22
27	20	0.5	40	3	US-09-306-290-41

28 20 0.5 649 4 US-09-569-751-238 Sequence 238, App  
29 20 0.5 706 4 US-09-702-705-164 Sequence 164, App  
30 20 0.5 706 4 US-09-736-457-164 Sequence 164, App  
31 20 0.5 771 3 US-09-040-984-85 Sequence 85, Appl  
32 20 0.5 771 4 US-09-123-912-85 Sequence 85, Appl  
33 20 0.5 771 4 US-09-643-597-85 Sequence 85, Appl  
34 20 0.5 771 4 US-09-480-884A-85 Sequence 85, Appl  
35 20 0.5 771 4 US-09-542-615A-85 Sequence 85, Appl  
36 20 0.5 771 4 US-09-606-421B-85 Sequence 85, Appl  
37 20 0.5 855 6 5185441-40 Patent No. 5185441  
38 20 0.5 855 6 5223394-3 Patent No. 5223394  
39 20 0.5 863 1 US-07-940-861-11 Sequence 11, Appl  
40 20 0.5 863 1 US-08-459-512-11 Sequence 11, Appl  
41 20 0.5 863 2 US-08-459-657-11 Sequence 11, Appl  
42 20 0.5 863 2 US-08-460-132-11 Sequence 11, Appl  
43 20 0.5 863 5 PCT-US92-02050-11 Patent No. 5185441  
44 20 0.5 863 6 5185441-35 Patent No. 5185441  
45 20 0.5 863 6 5223394-5 Patent No. 5223394

#### ALIGNMENTS

RESULT 1  
US-09-620-312D-45  
; Sequence 45, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yungqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: NO. 6569662e1 Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620.312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: dc-fl\_genes Version 1.0  
; SEQ ID NO 45  
; LENGTH: 2905  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (173)..(2452)  
US-09-620-312D-45

Query Match 69.2%; Score 2565; DB 4; Length 2905;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 22 TAGCAGCTTGGTCCGACAGGTCGGTAGTAGAGCGCGGACCTGTACAGGGCTGTT 81  
Db 58 TAGCAGCTTGGTCCGACAGGTCGGTAGTAGAGCGCGGACCTGTACAGGGCTGTT 117

QY 82 AGCAGCGCAGAGAAAGGGCGCTTTTAGCCAGGTATTTTCAGTGTCTGTAGACAAGATGGA 141  
Db 118 AGCAGCGCAGAGAAAGGGCGCTTTTAGCCAGGTATTTTCAGTGTCTGTAGACAAGATGGA 177  
QY 142 ATCATCTCATTTAATAGACGGCAATGGACCTCACTATCATTTGAGGGTAAACGCCAAAGA 201  
Db 178 ATCATCTCATTTAATAGACGGCAATGGACCTCACTATCATTTGAGGGTAAACGCCAAAGA 237  
QY 202 ACTTTCTCTTGTCAACAAGAAAGTCATCGGCTATTGTGGAAATATTCTCCAAGTACCA 261  
Db 238 ACTTTCTCTTGTCAACAAGAAAGTCATCGGCTATTGTGGAAATATTCTCCAAGTACCA 297  
QY 262 GAAAGCAGCTGAAGAAACAAACATGGAGAAAGAGAGATTAACCGGAAATCTCTCCCA 321  
Db 298 GAAAGCAGCTGAAGAAACAAACATGGAGAAAGAGAGATTAACCGGAAATCTCTCCCA 357  
QY 322 GCNCTTTAGAAAGGGACCCCTGACTGTGTGTAAAGAAAGTGGAGAACCCAGGGCTGGG 381  
Db 358 GCACITTTAGAAAGGGACCCCTGACTGTGTGTAAAGAAAGTGGAGAACCCAGGGCTGGG 417  
QY 382 AGCAGAGTCTCACACAGACTCTACGGAAACAGCAGCACTAGATTAGCAGAGCAGA 441  
Db 418 AGCAGAGTCTCACACAGACTCTACGGAAACAGCAGCACTAGATTAGCAGAGCAGA 477  
QY 442 CCATCTCTCTGCTGAAGTGACAAAGCCACGCTCTTCTGAGGCCAAAGCTGACCCAGAGA 501  
Db 478 CCATCTCTCTGCTGAAGTGACAAAGCCACGCTCTTCTGAGGCCAAAGCTGACCCAGAGA 537  
QY 502 ACAATATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAGCCCTCGTTTCAGGGTGCATA 561  
Db 538 ACAATATCCACCCAGATCTAGACTCAGGTCACTCTCTGAAGCCCTCGTTTCAGGGTGCATA 597  
QY 562 TCCCCACATCAAGGACGCTGAGGATCTTAAGACCACTCAACAGAAATGTAATAAATGGA 621  
Db 598 TCCCCACATCAAGGACGCTGAGGATCTTAAGACCACTCAACAGAAATGTAATAAATGGA 657  
QY 622 AAATTTGTCTAGAGAAATCAGGATGAAGTGAAGGCTTAAGATGATGTTTGA 741  
Db 658 AAATTTGTCTAGAGAAATCAGGATGAAGTGAAGGCTTAAGATGATGTTTGA 777  
QY 742 GAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAAGCCGCAAGTGGGAAG 801  
Db 778 GAAAGGTGAACCAACTCAAACTAAGATTCTCCGGGCCCAAGCCGCAAGTGGGAAG 837  
QY 802 GAAGATCTCTCAAAACAGCTATTCTTAGATGACTGGAAATAGGCCCAAGTCAAGTGGAG 861  
Db 838 GAAGATCTCTCAAAACAGCTATTCTTAGATGACTGGAAATAGGCCCAAGTCAAGTGGAG 897  
QY 862 ATCTTCTACATTTGACTCGGAGAAATAGAGTAGACGAAATCTGGAACTTCCAGCCT 921  
Db 898 ATCTTCTACATTTGACTCGGAGAAATAGAGTAGACGAAATCTGGAACTTCCAGCCT 957  
QY 922 CTCGAAACCTCTATAAAGATCGAATGGCCCAAGTACCAGGACGCTGTGTCCAAACAAAG 981  
Db 958 CTCGAAACCTCTATAAAGATCGAATGGCCCAAGTACCAGGACGCTGTGTCCAAACAAAG 1017  
QY 982 CAGCTCAACCAACTATACAAATGAGCTGAAGCCAGTGGTGGGGAATCAAAATTCATAA 1041  
Db 1018 CAGCTCAACCAACTATACAAATGAGCTGAAGCCAGTGGTGGGGAATCAAAATTCATAA 1077  
QY 1042 AATCGGACAAAGAGAAATGTGCCCCAGGCTCTGAGGCTCTGATCACCCTCAGGAAG 1101  
Db 1078 AATCGGACAAAGAGAAATGTGCCCCAGGCTCTGAGGCTCTGATCACCCTCAGGAAG 1137  
QY 1102 GGAAGAAATTTCTGCAATAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCCGAGATGA 1161  
Db 1138 GGAAGAAATTTCTGCAATAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCCGAGATGA 1197  
QY 1162 CTCCCGTGAATCCAGGTTAAGAGTAGGTTCAACAGCCTGTCCATCCCAAGCCATAAG 1221

Db 1198 TCCCGTGACTCCCGAGTTAAGAGTGAGGTTCAACAGCCTGTCCATCCCAAGCCACTAAG 1257  
QY 1222 TCCAGATTTCCAGAGCCTCCAGTCTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAAGTT 1281  
Db 1258 TCCAGATTTCCAGAGCCTCCAGTCTTTCTGAAAGTCTCTCCCAAGCAATGAAGAAGTT 1317  
QY 1282 TCAGGCACCTCTCAAGAGAGACCTCGGTGGAATGTCAAGAGACAGTCTATCCAAATGGAGCG 1341  
Db 1318 TCAGGCACCTCTCAAGAGAGACCTCGGTGGAATGTCAAGAGACAGTCTATCCAAATGGAGCG 1377  
QY 1342 TCTCTTGGCCAAACAGCAGGTGTTTTCACATCAGCTGCTTCCGTTGCTCTATTGCAACAA 1401  
Db 1378 TCTCTTGGCCAAACAGCAGGTGTTTTCACATCAGCTGCTTCCGTTGCTCTATTGCAACAA 1437  
QY 1402 CAATCTCAGTCTAGGAACATATGCATCTTTTACATGGAAGAAATCTATTGTAAGCCTCACTT 1461  
Db 1438 CAATCTCAGTCTAGGAACATATGCATCTTTTACATGGAAGAAATCTATTGTAAGCCTCACTT 1497  
QY 1462 CAATCTCACTCTTTAAATCTTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACACACAA 1521  
Db 1498 CAATCTCACTCTTTAAATCTTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACACACAA 1557  
QY 1522 GGATCTTATGGGCAAGCAAAAATGAAACGAAAGAGATTTTGGAGAGACCAAGCCTGTC 1581  
Db 1558 GGATCTTATGGGCAAGCAAAAATGAAACGAAAGAGATTTTGGAGAGACCAAGCCTGTC 1617  
QY 1582 AAATGCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCCCTATTGCTAAGTGGG 1641  
Db 1618 AAATGCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCCCTATTGCTAAGTGGG 1677  
QY 1642 TGTCTCTGGCTGCAAGTATGGAAGCCAGGCTCTCTCAGCAGGAGAAAGAACAGAGCC 1701  
Db 1678 TGTCTCTGGCTGCAAGTATGGAAGCCAGGCTCTCTCAGCAGGAGAAAGAACAGAGCC 1737  
QY 1702 AGCTGAAACCAAGAAAGCTGAGGATCCCTGGCCACCCCACTGAACTTGGAAAGTTCAGG 1761  
Db 1738 AGCTGAAACCAAGAAAGCTGAGGATCCCTGGCCACCCCACTGAACTTGGAAAGTTCAGG 1797  
QY 1762 AAGTGCCTTGGAGGAAGGATCAAAATGTCAAGGCCCAATGGCCTCTGAAAGACGAAAT 1821  
Db 1798 AAGTGCCTTGGAGGAAGGATCAAAATGTCAAGGCCCAATGGCCTCTGAAAGACGAAAT 1857  
QY 1822 CAGCAAGCCGGAAGTCTCTGAGGATGTTCATCTAGATCTGAAGAAGCTTAAGACGATCTTC 1881  
Db 1858 CAGCAAGCCGGAAGTCTCTGAGGATGTTCATCTAGATCTGAAGAAGCTTAAGACGATCTTC 1917  
QY 1882 TTTCACTGAAAGAAAGAGCCGCCCATTTCACTGACAGCTTCATTTCAAGCACCTCTGT 1941  
Db 1918 TTTCACTGAAAGAAAGAGCCGCCCATTTCACTGACAGCTTCATTTCAAGCACCTCTGT 1977  
QY 1942 CAAGAGCCCAAAACTGTGTCTCCCACTATCAGGAAAGGCTGGAGCATGTTCAGAGCAGAG 2001  
Db 1978 CAAGAGCCCAAAACTGTGTCTCCCACTATCAGGAAAGGCTGGAGCATGTTCAGAGCAGAG 2037  
QY 2002 TGAAGAGTCTGTGGGTGGAAGGTTGCAAGAAAGGAAACAAAGTGGAAATGCAAGGCTTC 2061  
Db 2038 TGAAGAGTCTGTGGGTGGAAGGTTGCAAGAAAGGAAACAAAGTGGAAATGCAAGGCTTC 2097  
QY 2062 TAAAGAAATGGAATGTGGGAAACAAACCTGCGCAAAACAAAGAAATCTAAAGGAGAGAC 2121  
Db 2098 TAAAGAAATGGAATGTGGGAAACAAACCTGCGCAAAACAAAGAAATCTAAAGGAGAGAC 2157  
QY 2122 AGGGAAGAAAGTAAAGAGGCTCATAGTTTGGAGATGGAGAAATGAGAAATCTTTAGAAA 2181  
Db 2158 AGGGAAGAAAGTAAAGAGGCTCATAGTTTGGAGATGGAGAAATGAGAAATCTTTAGAAA 2217  
QY 2182 TGGTGCAGACTCCGATGAAGATGAACAGCTTCTCTCAACAAACAAATCTCCACAAAGACC 2241  
Db 2218 TGGTGCAGACTCCGATGAAGATGAACAGCTTCTCTCAACAAACAAATCTCCACAAAGACC 2277  
QY 2242 CAAGTCTCTGAATTTGGTCGAGTTTGTAGACAAACCTTTTGTCTGAAAGAAATTCCTACTCA 2301

```
Db 2278 CAAAGTCCTGAAATTTGGTCAGAGTTTGTAGACAACACCTTTTCTGTAAGAATTCACACTCTCA 2337
Qy 2302 GAATCAGAAATCCAGGATGTTGGAATCTGGGAGGAGAGTGGTCAAGAGCTCTCTGT 2361
Db 2338 GAATCAGAAATCCAGGATGTTGGAATCTGGGAGGAGAGTGGTCAAGAGCTCTCTGT 2397
Qy 2362 GGAAGACAGATAAAGAGAAATCGGTATTTATGATGAGGATGAGGATGAAGAGTGCACAAAT 2421
Db 2398 GGAAGACAGATAAAGAGAAATCGGTATTTATGATGAGGATGAGGATGAAGAGTGCACAAAT 2457
Qy 2422 TCGAATGATGCTGGGCCCTTAATTTTCATGTTAGTGTAGGAGCCACTGCCCTTTGTCAAA 2481
Db 2458 TCGAATGATGCTGGGCCCTTAATTTTCATGTTAGTGTAGGAGCCACTGCCCTTTGTCAAA 2517
Qy 2482 ATGTGATGCACATAAGCAGGATCCAGCATGAAATGTAATTTACTTGGAGTAACTTTG 2541
Db 2518 ATGTGATGCACATAAGCAGGATCCAGCATGAAATGTAATTTACTTGGAGTAACTTTG 2577
Qy 2542 GAAAGAAATTCCTTCTTAAATCAAAAAACAAAAACAAAAACACA 2586
Db 2578 GAAAGAAATTCCTTCTTAAATCAAAAAACAAAAACAAAAACACA 2622

RESULT 2
US-09-385-982-139/c
; Sequence 139, Application US/09385982
; Patent No. 6262334
; GENERAL INFORMATION:
; APPLICANT: ENDIGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCNA-260XX
; CURRENT APPLICATION NUMBER: US/09/385.982
; EARLIER FILING DATE: 1999-08-30
; EARLIER FILING DATE: 1999-06-08
; EARLIER FILING DATE: 1999-01-27
; EARLIER FILING DATE: 1999-01-27
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 139
; LENGTH: 586
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(586)
; OTHER INFORMATION: n = A, T, C or G
US-09-385-982-139

Query Match
Best Local Similarity 9.8%; Score 362; DB 3; Length 586;
Matches 362; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3137 GTACTCCAAATGATGCTTTCTTTCTGGTGATCTGTGCTTCTCATTAATTAAGGAG 3196
Db 363 GTACTCCAAATGATGCTTTCTTTCTGGTGATCTGTGCTTCTCATTAATTAAGGAG 304

Qy 3197 CTGCAATATTTAGTAATACCTTCGGGATCACTGTCCCCCACTTCCTGGTTAGAGCAA 3256
Db 303 CTGCAATATTTAGTAATACCTTCGGGATCACTGTCCCCCACTTCCTGGTTAGAGCAA 244

Qy 3257 GTGAAGAGTTTAAAGGAGGAAGAAAGAACTGCTTACACACCTTGAGCTCAGACCTC 3316
Db 243 GTGAAGAGTTTAAAGGAGGAAGAAAGAACTGCTTACACACCTTGAGCTCAGACCTC 184

Qy 3317 TAAACCTGTATTTCCCTTATGATGCCCTTTTGGAGACACTAAATTTTAAATTAATCTAC 3376
Db 183 TAAACCTGTATTTCCCTTATGATGCCCTTTTGGAGACACTAAATTTTAAATTAATCTAC 124

Qy 3377 TAGCTCTGAAATATATGATTTTATCAGACTATCTTCAGGGTGAAATTAACCAACTAT 3436
```

```
Db 123 TAGCTCTGAAATATATGATTTTATCAGAGTATCTCAGGGTGAATTTAAACCAACTAT 64
Qy 3437 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTATAAACTTCAGT 3496
Db 63 AGGCTTTTCTTGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTATAAACTTCAGT 4
Qy 3497 AC 3498
Db 3 AC 2

RESULT 3
US-09-016-434-493
; Sequence 493, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.434
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 493:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNO01
; CLONE: 2272243
US-09-016-434-493

Query Match
Best Local Similarity 7.4%; Score 276; DB 4; Length 279;
Matches 276; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2654 TAAAGTGTCTGAAGCTTGTAACTGGGGAATATTCACCTGATTAATAGCCAGATTCTAC 2713
Db 4 TAAAGTGTCTGAAGCTTGTAACTGGGGAATATTCACCTGATTAATAGCCAGATTCTAC 63

Qy 2714 TGTATTCCTCCAAAGGCAATATTAAGGTAGATAGTATGATGATTTGTTACACT 2773
Db 64 TGTATTCCTCCAAAGGCAATATTAAGGTAGATAGTATGATGATTTGTTACACT 123

Qy 2774 ATTTGGAATTAGAGAACATACAGAGGAATTTAGGGCTTAAACACTGACTGAATGC 2833
```

Db 124 ATTTTGGAAATTAGAGACATACAGAGAAATTTAGGGCTTAAACATTTACACTGAATGC 183  
QY 2834 ACTTTAGTATAAAGGCGCAGCTTTCTATATTTTAAATGAATACCAATTTAAATTTTAG 2893  
Db 184 ACTTTAGTATAAAGGCGCAGCTTTCTATATTTTAAATGAATACCAATTTAAATTTTAG 243  
QY 2894 TATTTACTGTGTAAGAGATTTATTTAGTCTTTAAATTT 2929  
Db 244 TATTTACTGTGTAAGAGATTTATTTAGTCTTTAAATTT 279

RESULT 4  
US-09-016-434-680  
; Sequence 680, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREMITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 680:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 413 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: HNT2NOT01  
; CLONE: 269891  
US-09-016-434-680

Query Match 6.7%; Score 248; DB 4; Length 413;  
Best Local Similarity 100.0%; Pred. No. 4.5e-97;  
Matches 248; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1183 GAGTGAGTTCAACAGCCTGTCCATCCCAAGCACTAAGTCCAGATTCAGAGCCTCCAG 1242  
Db 166 GAGTGAGTTCAACAGCCTGTCCATCCCAAGCACTAAGTCCAGATTCAGAGCCTCCAG 225  
QY 1243 TCTTCTGAAAGTTCTCTCCCAAGCACTAAGTCCAGATTCAGAGCCTCCAGAGAGAC 1302  
Db 226 TCTTCTGAAAGTTCTCTCCCAAGCACTAAGTCCAGATTCAGAGCCTCCAGAGAGAC 285  
QY 1303 CTGCGTGGATGTGAGAGACAGTCTATCCCAATGGAGCGTCTCTTGGCCCAACGAGGAT 1362

Db 286 CTGCGTGGATGTGAGAGACAGTCTATCCAAATGGAGCGTCTCTTGGCCCAACGAGGAT 345  
QY 1363 GTTTACATCAGCTGCTTCGGTTGCTCTTATTCGACAAACAACACTCAGTCTAGGACATA 1422  
Db 346 GTTTACATCAGCTGCTTCGGTTGCTCTTATTCGACAAACAACACTCAGTCTAGGACATA 405  
QY 1423 TGCATCTTT 1430  
Db 406 TGCATCTTT 413

RESULT 5  
US-09-016-434-181  
; Sequence 181, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Seilhamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,434  
; FILING DATE: HEREMITH  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 181:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 238 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: KIDNNOT09  
; CLONE: 1422786  
US-09-016-434-181

Query Match 3.2%; Score 119; DB 4; Length 238;  
Best Local Similarity 100.0%; Pred. No. 1.2e-41;  
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2208 ACAGTCTCTCTAAACAACAATCTCCACAAGAACCCCAAGTCTCTGAATTTGGTCGAGTTTG 2267  
Db 95 ACAGTCTCTCTAAACAACAATCTCCACAAGAACCCCAAGTCTCTGAATTTGGTCGAGTTTG 154  
QY 2268 TAGACAACACCTTTTGTGTAAGAAATTCATCTCTCAGAATCAGAAATCCCAAGGATGTGGAA 2326  
Db 155 TAGACAACACCTTTTGTGTAAGAAATTCATCTCTCAGAATCAGAAATCCCAAGGATGTGGAA 213

RESULT 6  
US-08-937-540-7  
; Sequence 7, Application US/08937540  
; Patent No. 5891697  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney B  
; APPLICANT: Wise, Mitchell L  
; APPLICANT: Savage, Thomas J  
; APPLICANT: Katschka, Eva J  
; TITLE OF INVENTION: Monoterpene Synthases from Common Sage  
; TITLE OF INVENTION: (Salvia officinalis)  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON & KINDNESS  
; STREET: 1420 FIFTH AVENUE  
; CITY: SEATTLE  
; STATE: WASHINGTON  
; COUNTRY: USA  
; ZIP: 98101-2347  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/937,540  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Shelton, Dennis K  
; REGISTRATION NUMBER: 26,997  
; REFERENCE/DOCKET NUMBER: WSU111254  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206 695 1718  
; TELEFAX: 206 224 0779  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2022 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cdna  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Salvia officinalis  
; IMMEDIATE SOURCE:  
; CLONE: Unknown monoterpene synthase-like sequence  
; US-08-937-540-7  
Query Match 0.6%; Score 24; DB 2; Length 2022;  
Best Local Similarity 100.0%; Pred.No. 0.69;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3682 TAAATTTGAAAAAAGAAAAA 3705  
Db 1974 TAAATTTGAAAAAAGAAAAA 1997  
RESULT 7  
US-07-867-106-4  
; Sequence 4, Application US/07867106  
; Patent No. 5389526  
; GENERAL INFORMATION:  
; APPLICANT: Slade, Martin B  
; APPLICANT: Chang, Andy C M  
; APPLICANT: Williams, Keith L  
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular  
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris  
; STREET: One Liberty Place 46th Floor

; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/867,106  
; FILING DATE: 19920625  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: AU PJ 7187  
; APPLICATION NUMBER: PCT/AU90/00530  
; FILING DATE: 02-NOV-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Feeney, Joanne Longo  
; REGISTRATION NUMBER: 35,134  
; REFERENCE/DOCKET NUMBER: RICE-0002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3138 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; US-07-867-106-4  
Query Match 0.6%; Score 24; DB 1; Length 3138;  
Best Local Similarity 100.0%; Pred.No. 0.67;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3681 TTAAATTTGAAAAAAGAAAAA 3704  
Db 2443 TTAAATTTGAAAAAAGAAAAA 2466  
RESULT 8  
US-08-330-108-4  
; Sequence 4, Application US/08330108  
; Patent No. 5795752  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Kendall A.  
; TITLE OF INVENTION: IL-2-Stimulated Gene  
; TITLE OF INVENTION: Expression  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Lahive & Cockfield  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/330,108  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/104,736  
; FILING DATE:  
; APPLICATION NUMBER: US/07/796,066  
; FILING DATE:  
; INFORMATION FOR SEQ ID NO: 4:



```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 199 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; CELL TYPE: T-cell blast
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: 10A8
; US-08-330-108-4

Query Match 0.6%; Score 23; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3683 AAAATTGAAAAA 3705
Db 177 AAAATTGAAAAA 199

RESULT 10
US-09-345-468-1
; Sequence 1, Application US/09345468
; Patent No. 624527
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/345,468
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-345-468-1

Query Match 0.6%; Score 23; DB 3; Length 2047;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2568 AACAAAAACAAAAACACAAAAA 2590
Db 1683 AACAAAAACAAAAACACAAAAA 1705

RESULT 11
US-09-414-453A-1
; Sequence 1, Application US/09414453A
; Patent No. 6383779
; GENERAL INFORMATION:
; APPLICANT: Busfield, S.
; APPLICANT: Villevall, J.
; APPLICANT: Jandrot-Perrus, M.
; APPLICANT: Vainchenker, W.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-147
; CURRENT APPLICATION NUMBER: US/09/414,453A
; CURRENT FILING DATE: 1999-10-07
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-414-453A-1

Query Match 0.6%; Score 23; DB 4; Length 2047;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2568 AACAAAAACAAAAACACAAAAA 2590
Db 1683 AACAAAAACAAAAACACAAAAA 1705

; SEQUENCE CHARACTERISTICS:
; LENGTH: 199 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; CELL TYPE: T-cell blast
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: 10A8
; US-08-330-108-4

Query Match 0.6%; Score 23; DB 1; Length 199;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3683 AAAATTGAAAAA 3705
Db 177 AAAATTGAAAAA 199

RESULT 9
PCT-US92-10087-4
; Sequence 4, Application PC/TUS9210087
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A.
; TITLE OF INVENTION: IL-2-Stimulated Gene
; TITLE OF INVENTION: Expression
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lahive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10087
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/796,066
; FILING DATE: 20-NOVEMBER-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-028PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 199 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; CELL TYPE: T-cell blast
; IMMEDIATE SOURCE:
; LIBRARY:
```

```
RESULT 12
US-09-832-312-1
; Sequence 1, Application US/09832312
; Patent No. 6548741
; GENERAL INFORMATION:
; APPLICANT: Busfield et al.
; TITLE OF INVENTION: GLYCOPROTEIN VI AND USES THEREOF
; FILE REFERENCE: 7853-234
; CURRENT APPLICATION NUMBER: US/09/832,312
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/610,118
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/503,387
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: 09/454,824
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: 09/345,468
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-832-312-1

Query Match      0.6%; Score 23; DB 4; Length 2047;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2568 AACAAACACAAAAACACAAAAA 2590
DB 1683 AACAAACACAAAAACACAAAAA 1705

RESULT 13
US-09-276-531-8/c
; Sequence 8, Application US/09276531
; Patent No. 6183968
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lai, Freeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Reddy, Roopa
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF GENES ENCODING
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/276,531
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/079,677
; FILING DATE: March 27, 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lynn E. Murry, Ph.D.

US-09-276-531-8
; REGISTRATION NUMBER: 42,918
; REFERENCE/DOCKET NUMBER: PA-0008 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3556 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: LIVRUT01
; CLONE: 1753826
US-09-276-531-8

Query Match      0.6%; Score 23; DB 3; Length 3556;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3682 TAAATTTGAAAAAAGAAAAA 3704
DB 2982 TAAATTTGAAAAAAGAAAAA 2960

RESULT 14
US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSET-031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
```

OTHER INFORMATION:	5-130-257	:	polymorphic base A or G
FEATURE:			
NAME/KEY:	allele		
LOCATION:	99117		
OTHER INFORMATION:	5-130-276	:	polymorphic base A or G
FEATURE:			
NAME/KEY:	allele		
LOCATION:	103806		
OTHER INFORMATION:	5-131-395	:	polymorphic base A or T
FEATURE:			
NAME/KEY:	allele		
LOCATION:	106940		
OTHER INFORMATION:	5-133-375	:	polymorphic base insertion of A
FEATURE:			
NAME/KEY:	allele		
LOCATION:	108106		
OTHER INFORMATION:	5-135-155	:	polymorphic base insertion of A
FEATURE:			
NAME/KEY:	allele		
LOCATION:	108149		
OTHER INFORMATION:	5-135-198	:	polymorphic base insertion of GTTT
FEATURE:			
NAME/KEY:	allele		
LOCATION:	108308		
OTHER INFORMATION:	5-135-357	:	polymorphic base A or G
FEATURE:			
NAME/KEY:	allele		
LOCATION:	108471		
OTHER INFORMATION:	5-136-174	:	polymorphic base C or T
FEATURE:			
NAME/KEY:	allele		
LOCATION:	134134		
OTHER INFORMATION:	5-140-120	:	polymorphic base C or T
FEATURE:			
NAME/KEY:	allele		
LOCATION:	134362		
OTHER INFORMATION:	5-140-348	:	polymorphic base insertion of A
FEATURE:			
NAME/KEY:	allele		
LOCATION:	134374		
OTHER INFORMATION:	5-140-361	:	polymorphic base insertion of CA
FEATURE:			
NAME/KEY:	allele		
LOCATION:	146328		
OTHER INFORMATION:	5-143-84	:	polymorphic base A or G
FEATURE:			
NAME/KEY:	allele		
LOCATION:	146345		
OTHER INFORMATION:	5-143-101	:	polymorphic base A or C
FEATURE:			
NAME/KEY:	allele		
LOCATION:	150329		
OTHER INFORMATION:	5-145-24	:	polymorphic base A or G
FEATURE:			
NAME/KEY:	allele		
LOCATION:	160031		
OTHER INFORMATION:	5-148-352	:	polymorphic base G or T
FEATURE:			
NAME/KEY:	allele		
LOCATION:	72771..72817		
OTHER INFORMATION:	polymorphic fragment 5-124-273	SEQ ID30	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	72771..72817		
OTHER INFORMATION:	polymorphic fragment 5-124-273	SEQ ID51	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	88050..88096		
OTHER INFORMATION:	polymorphic fragment 5-127-261	SEQ ID31	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	88050..88096		
OTHER INFORMATION:	polymorphic fragment 5-127-261	SEQ ID52	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	90819..90865		
OTHER INFORMATION:	complement polymorphic fragment 99-1437-325	SEQ ID49	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	90819..90865		
OTHER INFORMATION:	complement polymorphic fragment 99-1437-325	SEQ ID70	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	93690..93736		
OTHER INFORMATION:	polymorphic fragment 5-128-60	SEQ ID32	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	93690..93736		
OTHER INFORMATION:	polymorphic fragment 5-128-60	SEQ ID53	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	97099..97145		
OTHER INFORMATION:	polymorphic fragment 99-1442-224	SEQ ID50	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	97099..97145		
OTHER INFORMATION:	polymorphic fragment 99-1442-224	SEQ ID71	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	97130..97177		
OTHER INFORMATION:	polymorphic fragment 5-129-144	SEQ ID33	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	97130..97177		
OTHER INFORMATION:	polymorphic fragment 5-129-144	SEQ ID54	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	99075..99121		
OTHER INFORMATION:	polymorphic fragment 5-130-257	SEQ ID34	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	99075..99121		
OTHER INFORMATION:	polymorphic fragment 5-130-257	SEQ ID55	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	99094..99140		
OTHER INFORMATION:	polymorphic fragment 5-130-276	SEQ ID35	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	99094..99140		
OTHER INFORMATION:	polymorphic fragment 5-130-276	SEQ ID56	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	103783..103828		
OTHER INFORMATION:	polymorphic fragment 5-131-395	SEQ ID36	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	103783..103828		
OTHER INFORMATION:	polymorphic fragment 5-131-395	SEQ ID57	
FEATURE:			
NAME/KEY:	allele		
LOCATION:	106918..106966		
OTHER INFORMATION:	polymorphic fragment 5-133-375	SEQ ID37	
FEATURE:			

```
; NAME/KEY: allele
; LOCATION: 108127...108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108127...108177
; OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
; FEATURE:
```

```
Query Match          0.6%; Score 22; DB 4; Length 162450;
Best Local Similarity 100.0%; Pred.No. 3.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2564 CAAAACACAAACAAACAAACAC 2585
      |||||
DB 136286 CAAAACAAACAAACAAACAC 136265
```

## RESULT 15

```
US-08-064-694-1
; Sequence 1, Application US/08064694
; Patent No. 5847076
; GENERAL INFORMATION:
; APPLICANT: George N. DeMartino
; APPLICANT: Clive A. Slaughter
; APPLICANT: Patricia J. Willy
; APPLICANT: Ma Chu-Ping
; TITLE OF INVENTION: REGULATORS OF THE PROTEASOME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/064,694
; FILING DATE: 19930520
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Hodgins, Daniel S.
; REGISTRATION NUMBER: 31,026
; REFERENCE/DOCKET NUMBER: UTSD:356/HOD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512-320-7200
; TELEFAX: 512-474-7577
; TELEX: No. 5847076 Applicable
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-08-064-694-1
```

```
Query Match          0.6%; Score 21; DB 2; Length 720;
Best Local Similarity 100.0%; Pred.No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1679 CAGCAGAGAAGGAGACAAAG 1699
      |||||
DB 63 CAGCAGAGAAGGAGACAAAG 83
```

Search completed: January 6, 2004, 22:10:23  
Job time : 201 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 16:13:06 ; Search time 1105 Seconds

(without alignments)  
11605.454 Million cell updates/sec

Title: US-09-890-549-16

Perfect score: 3705

Sequence: 1 99cgcaggagcagtgagg9.....atttgaaaaa3705

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 2263443 seqs, 1730637950 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2589	69.9	2749	9	US-09-822-849A-399
2	2565	69.2	2905	13	US-10-117-722-45
3	2565	69.2	2905	15	US-10-037-270-45
4	2130	57.5	3650	11	US-09-783-732-3
5	2041	55.1	3543	11	US-09-783-732-1
6	556	15.0	732	9	US-09-925-297-152
7	423	11.4	547	15	US-10-066-543-3200
8	415	11.2	565	15	US-10-060-036-4338
9	401	10.8	583	15	US-10-066-543-3373
10	385	10.4	698	15	US-10-198-846-10895
11	362	9.8	586	11	US-09-871-161-139
12	358	9.7	358	15	US-10-066-543-1533
13	253	6.8	262	10	US-09-796-692-2900
14	253	6.8	262	15	US-10-040-862-2900
15	212	5.7	567	15	US-10-198-846-8531

C 16	186	5.0	283	10	US-09-998-598-1767	Sequence 1767, Ap
C 17	175	4.7	175	10	US-09-878-178-2086	Sequence 2086, Ap
C 18	175	4.7	175	14	US-10-046-935-2086	Sequence 2086, Ap
C 19	175	4.7	175	15	US-10-146-502-2086	Sequence 2086, Ap
C 20	145	3.9	404	11	US-09-918-995-7225	Sequence 7225, Ap
C 21	140	3.8	518	15	US-10-198-846-1527	Sequence 1527, Ap
C 22	116	3.1	574	13	US-10-027-632-25035	Sequence 25035, A
C 23	116	3.1	574	14	US-10-027-632-25035	Sequence 25035, A
C 24	86	2.3	86	9	US-09-777-564-1052	Sequence 1052, Ap
C 25	86	2.3	86	15	US-10-015-219-1052	Sequence 1052, Ap
C 26	78	2.1	410	11	US-09-918-995-7296	Sequence 7296, Ap
C 27	73	2.0	228	10	US-09-783-590-176	Sequence 176, App
C 28	60	1.6	60	13	US-09-908-975-7077	Sequence 7077, Ap
C 29	42	1.1	840	13	US-09-865-879-3	Sequence 3, Appli
C 30	27	0.7	6334	13	US-10-311-455-1185	Sequence 1185, Ap
C 31	26	0.7	12968	13	US-10-311-455-2057	Sequence 2057, Ap
C 32	26	0.7	12968	13	US-10-240-453-297	Sequence 297, App
C 33	26	0.7	12968	15	US-10-239-676-201	Sequence 201, App
C 34	25	0.7	488	9	US-09-864-761-10140	Sequence 10140, A
C 35	25	0.7	1561	9	US-09-925-302-285	Sequence 265, App
C 36	24	0.6	552	13	US-10-027-632-248895	Sequence 248895,
C 37	24	0.6	552	14	US-10-027-632-248895	Sequence 248895,
C 38	24	0.6	793	10	US-09-878-574-4304	Sequence 4304, Ap
C 39	24	0.6	918	15	US-10-007-280A-114	Sequence 114, App
C 40	24	0.6	1593	11	US-09-764-891-9700	Sequence 9700, Ap
C 41	24	0.6	1593	15	US-10-205-428-947	Sequence 947, App
C 42	24	0.6	2753	15	US-10-007-280A-115	Sequence 115, App
C 43	24	0.6	7445	13	US-10-311-455-823	Sequence 823, App
C 44	24	0.6	17527	13	US-10-311-455-1406	Sequence 1406, Ap
C 45	23	0.6	380	15	US-10-102-524-1272	Sequence 1272, Ap

ALIGNMENTS

RESULT 1

US-09-822-849A-399  
; Sequence 399, Application US/09822849A  
; Patent No. US20020045170A1  
; GENERAL INFORMATION:  
; APPLICANT: Wong, Gordon G.  
; APPLICANT: Clark, Hilary  
; APPLICANT: Fechtel, Kim  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Resnick, Richard J.  
; APPLICANT: Gulukota, Kamalakr  
; APPLICANT: Graham, James R.  
; APPLICANT: Genetics Institute, Inc  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS  
; FILE REFERENCE: GIN 6403  
; CURRENT APPLICATION NUMBER: US/09/822, 849A  
; CURRENT FILING DATE: 2001-09-04  
; PRIOR APPLICATION NUMBER: 60/195,582  
; PRIOR FILING DATE: 2000-04-06  
; NUMBER OF SEQ ID NOS: 598  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 399  
; LENGTH: 2749  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-822-849A-399

Query Match	69.9%	Score 2589;	DB 9;	Length 2749;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2639;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
OY	1051	AAAGAGAAATGCCCCAGGTCCTGAGTCTGCCATCACCATCAGGAGGGAAGAT	1110	
Db	110	AAAGAGAAATGCCCCAGGTCCTGAGTCTGCCATCACCATCAGGAGGGAAGAT	169	
OY	1111	TTCTCAATGAGATACCTGGAGTCGTTCCACCCCTCCGGAAGATGATCTCCGTGA	1170	

170	Db	TTCTGCAAAATGAGATAGAGCTGGCAGTCCGTTCCACCCCTGCGGAAGATGACTCCCGTGA	229
1171	QY	CTCCAGGTTAAGAGTGAAGTTCAACAGAGCTGTCCATCCCAAGCCACTAAGTCCAGATTC	1230
230	Db	CTCCAGGTTAAGAGTGAAGTTCAACAGAGCTGTCCATCCCAAGCCACTAAGTCCAGATTC	289
1231	QY	CAGAGCCTCAGTCTTTCTGAAAGTTCTCTGCCAAAGCAATGAAGAAGTTTCAGGCACC	1290
290	Db	CAGAGCCTCAGTCTTTCTGAAAGTTCTCTGCCAAAGCAATGAAGAAGTTTCAGGCACC	349
1291	QY	TGCAAGAGAGACCTGCGTGGAAATGTCAGAAAGCAGTCTATCCAATGGAGCGTCTCTTGGC	1350
350	Db	TGCAAGAGAGACCTGCGTGGAAATGTCAGAAAGCAGTCTATCCAATGGAGCGTCTCTTGGC	409
1351	QY	CAACCAGCAGGTGTTTCACATCAGCTGCTTCCGTTGCTCTCTATTGCAACAACAACTCAG	1410
410	Db	CAACCAGCAGGTGTTTCACATCAGCTGCTTCCGTTGCTCTCTATTGCAACAACAACTCAG	469
1411	QY	TCTAGGAACATATGCAATCTTTTACATGGAAGATCTATTTGAAGCCTCACTTCAATCAACT	1470
470	Db	TCTAGGAACATATGCAATCTTTTACATGGAAGATCTATTTGAAGCCTCACTTCAATCAACT	529
1471	QY	CTTTAAATCTAAGGGCAACTATCATGAAGCTTTTGGGCACAGACACACAAGGATCTATG	1530
530	Db	CTTTAAATCTAAGGGCAACTATCATGAAGCTTTTGGGCACAGACACACAAGGATCTATG	589
1531	QY	GGCAAGCAAAAATGAAAAAGAGATTTTGGAGACAGCCAGCTTGCAAAATGCAAG	1590
590	Db	GGCAAGCAAAAATGAAAAAGAGATTTTGGAGACAGCCAGCTTGCAAAATGCAAG	649
1591	QY	GGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTATTGCTAAGGTGGGTGTCTTGGC	1650
650	Db	GGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTATTGCTAAGGTGGGTGTCTTGGC	709
1651	QY	TGCAAGTATGGAAGCCAAAGCCCTCTCTCAGCAGGAGAGAGAGCAAGCCAGCTGAAAC	1710
710	Db	TGCAAGTATGGAAGCCAAAGCCCTCTCTCAGCAGGAGAGAGAGCAAGCCAGCTGAAAC	769
1711	QY	CAAGAAGCTGAGGATCGCTTGGCCACCCCCCACTGAACTTTGGAAGTTTCAGGAAGTGCCTT	1770
770	Db	CAAGAAGCTGAGGATCGCTTGGCCACCCCCCACTGAACTTTGGAAGTTTCAGGAAGTGCCTT	829
1771	QY	GGAGGAAGGGATCAAAATGTCAAAGGCCAAATGGCCTCTGGAAGAGAAATCAGCAAGCC	1830
830	Db	GGAGGAAGGGATCAAAATGTCAAAGGCCAAATGGCCTCTGGAAGAGAAATCAGCAAGCC	889
1831	QY	CGAAGTCTCTGAGGATGTGCATCTAGATCTGAAGAAGCTAAGACGATCTTCTTCACTGAA	1890
890	Db	CGAAGTCTCTGAGGATGTGCATCTAGATCTGAAGAAGCTAAGACGATCTTCTTCACTGAA	949
1891	QY	GGAAAGAGCGGCCATTCTACTGTAGCAGCTTCATTTCAAAGCACCTCTCTCAAGAGCCC	1950
950	Db	GGAAAGAGCGGCCATTCTACTGTAGCAGCTTCATTTCAAAGCACCTCTCTCAAGAGCCC	1009
1951	QY	AAAAACTGTGTCCCCACCTATCAGGAAGGCTGGAGCATCTCAGAGCAGAGTGAAGATC	2010
1010	Db	AAAAACTGTGTCCCCACCTATCAGGAAGGCTGGAGCATCTCAGAGCAGAGTGAAGATC	1069
2011	QY	TGTGGGTGGAAGAGTTGCAGAAAGGAACAAGTGGAAAAATGCCAGGCTTCTTAAGAGAA	2070
1070	Db	TGTGGGTGGAAGAGTTGCAGAAAGGAACAAGTGGAAAAATGCCAGGCTTCTTAAGAGAA	1129
2071	QY	TGGGAATGTCGGAAGAAAACACCTTGGCAAAAACAAGATCTTAAAGGAGAGACGGGAGAG	2130
1130	Db	TGGGAATGTCGGAAGAAAACACCTTGGCAAAAACAAGATCTTAAAGGAGAGACGGGAGAG	1189
2131	QY	AAGTAAAGGAAGGTCATAGTTTGGAGATGGAAGATGAGAATCTTTGTAGAAAATGGTGCAGA	2190
1190	Db	AAGTAAAGGAAGGTCATAGTTTGGAGATGGAAGATGAGAATCTTTGTAGAAAATGGTGCAGA	1249
2191	QY	CTCCGATGAAGATGATTAACAGCTTCTTCAAACAACAAATCTTCCACAAGAACCCAGTCTCT	2250
1250	Db	CTCCGATGAAGATGATTAACAGCTTCTTCAAACAACAAATCTTCCACAAGAACCCAGTCTCT	1309

QY	2251	GAATTGGTCGAGTTTTTGTAGACAAACACCTTTTCTCTGAAGAAATTCACCTACTCTCAGAATTCAGAA	2310
Db	1310	GAATTGGTCGAGTTTTTGTAGACAAACACCTTTTCTCTGAAGAAATTCACCTACTCTCAGAATTCAGAA	1369
QY	2311	ATCCCAAGGATGTGGAACTCTGGGAGGGAAGTGGTCAAAGAGCTCTCTGTGGGAAGACA	2370
Db	1370	ATCCCAAGGATGTGGAACTCTGGGAGGGAAGTGGTCAAAGAGCTCTCTGTGGGAAGACA	1429
QY	2371	GATAAAGAGAAATTCGGTATTATGATGAGGATGAGGATGAAGAGTACAAAATTCGAATGAT	2430
Db	1430	GATAAAGAGAAATTCGGTATTATGATGAGGATGAGGATGAAGAGTACAAAATTCGAATGAT	1489
QY	2431	GCTGGGCCTTAAATTCATGTTTATGTTTACGAGGCCACTGCCCTTTTGTCTMAAATTTGTATGCC	2490
Db	1490	GCTGGGCCTTAAATTCATGTTTATGTTTACGAGGCCACTGCCCTTTTGTCTMAAATTTGTATGCC	1549
QY	2491	ACATAAGCAGGTATCCAGCATGAATAATGTAATTTACTTCGAAGTAACTTTTGGAAAGAAT	2550
Db	1550	ACATAAGCAGGTATCCAGCATGAATAATGTAATTTACTTCGAAGTAACTTTTGGAAAGAAT	1609
QY	2551	TCCTCTTAAATTCAAAAACAAAAACAAAAACAAAAACACATTTCTTAAATCTAGAG	2610
Db	1610	TCCTCTTAAATTCAAAAACAAAAACAAAAACAAAAACACATTTCTTAAATCTAGAG	1669
QY	2611	ATAACTTTTACTTAAATTCCTTCAATTTTAGCAGTGAATATGCATTAAGTCTCTTAAGGCTT	2670
Db	1670	ATAACTTTTACTTAAATTCCTTCAATTTTAGCAGTGAATATGCATTAAGTCTCTTAAGGCTT	1729
QY	2671	GTAACCTGGGAAATATTCACCTGATTAATAGCCAGATTTCTACTGTATTTCCCAAAAGGCA	2730
Db	1730	GTAACCTGGGAAATATTCACCTGATTAATAGCCAGATTTCTACTGTATTTCCCAAAAGGCA	1789
QY	2731	ATATTAAAGTGAATACATGATAGTAGTAGTATATTGTTACACACTATTTTCGAATTACAGAA	2790
Db	1790	ATATTAAAGTGAATACATGATAGTAGTAGTATATTGTTACACACTATTTTCGAATTACAGAA	1849
QY	2791	CATACAGAAGAAATTTAGGGGCTTTAAACATTCACGACTGAATGCACCTTTTGTATTAAGGGC	2850
Db	1850	CATACAGAAGAAATTTAGGGGCTTTAAACATTCACGACTGAATGCACCTTTTGTATTAAGGGC	1909
QY	2851	ACAGTTTGTATATTTTTAAATGAATACCAATTTAAATTTTTTACTGTGTTAAAGAG	2910
Db	1910	ACAGTTTGTATATTTTTAAATGAATACCAATTTAAATTTTTTACTGTGTTAAAGAG	1969
QY	2911	ATTATTTAGTCTTTAAATTTTTTAAAGTTAAATTTCTTGTCTGTATATATGAGGAATTT	2970
Db	1970	ATTATTTAGTCTTTAAATTTTTTAAAGTTAAATTTCTTGTCTGTATATATGAGGAATTT	2029
QY	2971	ACTACTTTATGTCCTGCTCTCTAAACTACATCCTGAACTCGACGCTCTCGAGGTATTAATAC	3030
Db	2030	ACTACTTTATGTCCTGCTCTCTAAACTACATCCTGAACTCGACGCTCTCGAGGTATTAATAC	2089
QY	3031	AACAGAGCACTTTTGGAGCAATTTGAAAAACCAACCTACACTCTTCGGTGCTTTAGAGAGA	3090
Db	2090	AACAGAGCACTTTTGGAGCAATTTGAAAAACCAACCTACACTCTTCGGTGCTTTAGAGAGA	2149
QY	3091	TCGTCTGTCTCCCAATAAGCTTTTGTATCTGSCAGTGAATTTACTGTACTCCAAATGAT	3150
Db	2150	TCGTCTGTCTCCCAATAAGCTTTTGTATCTGSCAGTGAATTTACTGTACTCCAAATGAT	2209
QY	3151	TGCTTTCTTTCTTGGTGATATCTGTCCTCTCATATAATTACTGAAAGCTCAATATTTTAG	3210
Db	2210	TGCTTTCTTTCTTGGTGATATCTGTCCTCTCATATAATTACTGAAAGCTCAATATTTTAG	2269
QY	3211	TAATACCTTCGGGATCACCTGCCCCCATCTTCGGTGTTTAGAGCAAAAGTGAAGAGTTTAAA	3270
Db	2270	TAATACCTTCGGGATCACCTGCCCCCATCTTCGGTGTTTAGAGCAAAAGTGAAGAGTTTAAA	2329
QY	3271	GGAGGAAGAAGAAGAACTGCTCTTACCACTTTGAGCTCAGACCTCTAAACCCCTGTATTT	3330
Db	2330	GGAGGAAGAAGAAGAACTGCTCTTACCACTTTGAGCTCAGACCTCTAAACCCCTGTATTT	2389

Qy 3331 CCCTTATGATGCCCTTTTGGAGACACTAAATTTTAAATACCTTACTAGCTCTGMAATAT 3390  
Db |||||  
Qy 2390 CCCTTATGATGCCCTTTTGGAGACACTAAATTTTAAATACCTTACTAGCTCTGMAATAT 2449  
Db |||||  
Qy 3391 ATTCAATTTTATCACAGTATTTCTCAGGGTGAATTTAAACCAACTATAGGCCCTTTTCTTG 3450  
Db |||||  
Qy 2450 ATTGAATTTTATCACAGTATTTCTCAGGGTGAATTTAAACCAACTATAGGCCCTTTTCTTG 2509  
Db |||||  
Qy 3451 GGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACCTAGTACATTTTGTGTACA 3510  
Db |||||  
Qy 2510 GGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACCTAGTACATTTTGTGTACA 2569  
Db |||||  
Qy 3511 CAGTTGATATCCAAATTTGATCGATGGGAGGAGGTTCTTAAAGCTGTAGGCTTTTC 3570  
Db |||||  
Qy 2570 CAGTTGATATCCAAATTTGATCGATGGGAGGAGGTTCTTAAAGCTGTAGGCTTTTC 2629  
Db |||||  
Qy 3571 TTTGTACTGCATTTATAGATTTAGCTTTAATTTTATAGATCTTAAACATTTCTGC 3630  
Db |||||  
Qy 2630 TTTGTACTGCATTTATAGATTTAGCTTTAATTTTATAGATCTTAAACATTTCTGC 2689  
Db |||||  
Qy 3631 TTTCTTAGTCTTACCTAGTCTGAAACATTTTATTCATTAAGATTTTAAATTTG 3690  
Db |||||  
Qy 2690 TTTCTTAGTCTTACCTAGTCTGAAACATTTTATTCATTAAGATTTTAAATTTG 2749  
Db |||||

RESULT 2

US-10-117-722-45  
; Sequence 45, Application US/10117722  
; Publication No. US20030219744A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219744A1ele Nucleic Acids and  
; FILE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BCIP  
; CURRENT APPLICATION NUMBER: US/10/117,722  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_Fl\_Genes version 1.0  
; SEQ ID NO 45  
; LENGTH: 2905  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (173)..(2452)  
US-10-117-722-45

Query Match 69.2%; Score 2565; DB 13; Length 2905;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 TAGCAGCTTGGTCGCGACAGGTGGCTTAGGTAGAGCGCGGACCTGTGACAGGGCTGGT 81  
Db |||||  
Qy 58 TAGCAGCTTGGTCGCGACAGGTGGCTTAGGTAGAGCGCGGACCTGTGACAGGGCTGGT 117  
Db |||||  
Qy 82 AGCAGCGCAGAGGAAAGCGCGCTTTTAGCCAGGTATTTACGTCTGTGTAGACAAGATGGA 141  
Db |||||  
Qy 118 AGCAGCGCAGAGGAAAGCGCGCTTTTAGCCAGGTATTTACGTCTGTGTAGACAAGATGGA 177  
Db |||||  
Qy 142 ATCATCTCCATTTAATAGACGGCAATGACCTCACTATCATTTAGGGTAAACAGCCAAAGA 201  
Db |||||  
Qy 178 ATCATCTCCATTTAATAGACGGCAATGACCTCACTATCATTTAGGGTAAACAGCCAAAGA 237  
Db |||||

Qy 202 ACTTCTCTTGTCAACAAGAACAAAGTCATCGCTATTGTGAAATATTTCTCCAAGTACCA 261  
Db |||||  
Qy 238 ACTTCTCTTGTCAACAAGAACAAAGTCATCGCTATTGTGAAATATTTCTCCAAGTACCA 297  
Db |||||  
Qy 262 GAAAGCAGCTGAAGAAACAAACATCGGAGAGAGAGAGTAACACCGGAAATCTCTCCCA 321  
Db |||||  
Qy 298 GAAAGCAGCTGAAGAAACAAACATCGGAGAGAGAGAGTAACACCGGAAATCTCTCCCA 357  
Db |||||  
Qy 322 GCATTTTGAAGGGGACCTTGACTGTGTTTAAAGAAAGTGGAGAACCCAGGGCTGGG 381  
Db |||||  
Qy 358 GCATTTTGAAGGGGACCTTGACTGTGTTTAAAGAAAGTGGAGAACCCAGGGCTGGG 417  
Db |||||  
Qy 382 AGCAGAGTCTCACACAGACTCTCTACGGAAACAGCAGACTCAGATTAGGCACAGAGCAGA 441  
Db |||||  
Qy 418 AGCAGAGTCTCACACAGACTCTCTACGGAAACAGCAGACTCAGATTAGGCACAGAGCAGA 477  
Db |||||  
Qy 442 CCATCCTCTCTGAAGTGACAAGCCACGCTGCTTCTGGAGCCAAAGCTGACCAAGAAAGA 501  
Db |||||  
Qy 478 CCATCCTCTCTGAAGTGACAAGCCACGCTGCTTCTGGAGCCAAAGCTGACCAAGAAAGA 537  
Db |||||  
Qy 502 ACAAAATCCACCCAGATCTAGACTCAGGTCACTCTGTAAGCCCTCGTTTCAGGGTCGATA 561  
Db |||||  
Qy 538 ACAAAATCCACCCAGATCTAGACTCAGGTCACTCTGTAAGCCCTCGTTTCAGGGTCGATA 597  
Db |||||  
Qy 562 TCCCCACATCAAGGACGCTGAGGATCTTTAAAGACCACTCAACAGAAATGAAAAAATGGA 621  
Db |||||  
Qy 598 TCCCCACATCAAGGACGCTGAGGATCTTTAAAGACCACTCAACAGAAATGAAAAAATGGA 657  
Db |||||  
Qy 622 AAATTGTCTAGGAGAAATCCAGGCATGAAGTAGAAAAATCAGAAATCAGTGAACACACAGA 681  
Db |||||  
Qy 658 AAATTGTCTAGGAGAAATCCAGGCATGAAGTAGAAAAATCAGAAATCAGTGAACACACAGA 717  
Db |||||  
Qy 682 TGCTTCGGGCAAAATAGAGAAATATAATTTCCGCTGAACAGGCTTAAGATGATGTTGA 741  
Db |||||  
Qy 718 TGCTTCGGGCAAAATAGAGAAATATAATTTCCGCTGAACAGGCTTAAGATGATGTTGA 777  
Db |||||  
Qy 742 GAAAGGTGAACCAACTCAAACTAAGATTTCTCGGCGCCCAAGCCGAGTGAAGTGAAG 801  
Db |||||  
Qy 778 GAAAGGTGAACCAACTCAAACTAAGATTTCTCGGCGCCCAAGCCGAGTGAAGTGAAG 837  
Db |||||  
Qy 802 GAAGATCTCTGAAACACAGCTATTCTCTAGATGACTGGAATAGGCCCCAGGTCAGTTGTC 861  
Db |||||  
Qy 838 GAAGATCTCTGAAACACAGCTATTCTCTAGATGACTGGAATAGGCCCCAGGTCAGTTGTC 897  
Db |||||  
Qy 862 ATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAACCTTCCACGGCT 921  
Db |||||  
Qy 898 ATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAACCTTCCACGGCT 957  
Db |||||  
Qy 922 CTCAGAAACCTCTATAAAGGATCGAATGGCCAAAGTACCAGGCAGCTGTGTCCAAACAAAG 981  
Db |||||  
Qy 958 CTCAGAAACCTCTATAAAGGATCGAATGGCCAAAGTACCAGGCAGCTGTGTCCAAACAAAG 1017  
Db |||||  
Qy 982 CAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGGTGGCGGAAATCAAAATTCATAA 1041  
Db |||||  
Qy 1018 CAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGGTGGCGGAAATCAAAATTCATAA 1077  
Db |||||  
Qy 1042 AATGAGCAAAAGGAGAAATGCCCCCAGGTCCTGAGGTCTGCATCACCCTCAGGAGG 1101  
Db |||||  
Qy 1078 AATGAGCAAAAGGAGAAATGCCCCCAGGTCCTGAGGTCTGCATCACCCTCAGGAGG 1137  
Db |||||  
Qy 1102 GGAAGAAATTTCTGCAAAATGAGATAGCTGGCAGTCGTTCCACCCCTGCGGAGATGA 1161  
Db |||||  
Qy 1138 GGAAGAAATTTCTGCAAAATGAGATAGCTGGCAGTCGTTCCACCCCTGCGGAGATGA 1197  
Db |||||  
Qy 1162 CTCCTGCTGACTCCAGGTTAAGAGTGAAGTTCAACAGGCTGTCCATCCCAAGCCACTAAG 1221  
Db |||||  
Qy 1198 CTCCTGCTGACTCCAGGTTAAGAGTGAAGTTCAACAGGCTGTCCATCCCAAGCCACTAAG 1257  
Db |||||  
Qy 1222 TCCAGATTCAGAGGCTCCAGTCTTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAGTT 1281  
Db |||||  
Qy 1258 TCCAGATTCAGAGGCTCCAGTCTTTTCTGAAAGTTCTCTCCCAAGCAATGAAGAGTT 1317  
Db |||||  
Qy 1282 TCAGGCACCTCGAAGAGAGACCTCGCTGGAAATGTTCAGNAGACAGTCTCTCAATGGAGCG 1341  
Db |||||

Db 1318 TCAGGCACCTGCAAGAGAGACCTGCGTGGAAATGTCAGAGACAGTCTATCAATGGAGCG 1377  
Qy 1342 TCTCTGGCCAAACAGCAGAGTGTTTCACATCAGCTGCTTCGCTGCTCTTATGCAACAA 1401  
Db 1378 TCTCTGGCCAAACAGCAGAGTGTTTCACATCAGCTGCTTCGCTGCTCTTATGCAACAA 1437  
Qy 1402 CAAACTCAGTCTAGGAACATATGCACTCTTACATGGAAGAAATCTATGTAAGCCTCACTT 1461  
Db 1438 CAAACTCAGTCTAGGAACATATGCACTCTTACATGGAAGAAATCTATGTAAGCCTCACTT 1497  
Qy 1462 CAACTCAACTCTTTAAATCTAAGGCAACTATGATGAAGGCTTTGGGCACAGCACACAA 1521  
Db 1498 CAACTCAACTCTTTAAATCTAAGGCAACTATGATGAAGGCTTTGGGCACAGCACACAA 1557  
Qy 1522 GGATCTATGGGCAAGCAAAATGAAACGAGAGATTTTGGAGAGACAGCCAGCTTGC 1581  
Db 1558 GGATCTATGGGCAAGCAAAATGAAACGAGAGATTTTGGAGAGACAGCCAGCTTGC 1617  
Qy 1582 AAATGCAAGGAGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTTATTTGTAAGGTGGG 1641  
Db 1618 AAATGCAAGGAGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTTATTTGTAAGGTGGG 1677  
Qy 1642 TGTCTGCTGCAAGTATGGAAGCAAGGCTCTCTCAGCAGGAGAGGAAGCAAGGCC 1701  
Db 1678 TGTCTGCTGCAAGTATGGAAGCAAGGCTCTCTCAGCAGGAGAGGAAGCAAGGCC 1737  
Qy 1702 AGCTGAAACCAAGAGCTGAGGATCGCTGCGCACCCCTCACTCACTGGAAGTTCAGG 1761  
Db 1738 AGCTGAAACCAAGAGCTGAGGATCGCTGCGCACCCCTCACTCACTGGAAGTTCAGG 1797  
Qy 1762 AAGTGCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAATGGCTCTGGAAGCAAGAAAT 1821  
Db 1798 AAGTGCCTTGGAGGAAGGATCAAAATGTCAAAGCCCAATGGCTCTGGAAGCAAGAAAT 1857  
Qy 1822 CAGCAAGCCCGAAGTCTCTGAGGATGTGATCTAGATCTGAAGAGCTTAAGCATCTTC 1881  
Db 1858 CAGCAAGCCCGAAGTCTCTGAGGATGTGATCTAGATCTGAAGAGCTTAAGCATCTTC 1917  
Qy 1882 TTCACTGAAGGAAGCAAGCCGCCATTCACCTGTAGCAGCTTCATTTCAAGCAGCTCTGT 1941  
Db 1918 TTCACTGAAGGAAGCAAGCCGCCATTCACCTGTAGCAGCTTCATTTCAAGCAGCTCTGT 1977  
Qy 1942 CAAGAGCCCAAAATCTGTGCTCCCACTTATCAGGAAGGCTGGAGCATGTGAGAGCAGAG 2001  
Db 1978 CAAGAGCCCAAAATCTGTGCTCCCACTTATCAGGAAGGCTGGAGCATGTGAGAGCAGAG 2037  
Qy 2002 TGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAACAAGTGGAAAAATGCCAGGCTTC 2061  
Db 2038 TGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAACAAGTGGAAAAATGCCAGGCTTC 2097  
Qy 2062 TAAGAGAAATGGAAATGTGGGAAAAACAACCTGGCAAAACAAGAAATCTAAGAGGAGAC 2121  
Db 2098 TAAGAGAAATGGAAATGTGGGAAAAACAACCTGGCAAAACAAGAAATCTAAGAGGAGAC 2157  
Qy 2122 AGGGAAGAGAACTAAGGAAGGTCATAGTTTGGAGATGGAGATGAGATCTTTGTAGAAAA 2181  
Db 2158 AGGGAAGAGAACTAAGGAAGGTCATAGTTTGGAGATGGAGATGAGATCTTTGTAGAAAA 2217  
Qy 2182 TGGTGCAGACTCCGATGAAGATGATAACAGCTTCTCAAACAACAAATCTCCACAGAACCC 2241  
Db 2218 TGGTGCAGACTCCGATGAAGATGATAACAGCTTCTCAAACAACAAATCTCCACAGAACCC 2277  
Qy 2242 CAACTCTCTGAATTTGGTCAAGTTTTGTAGAACAACTTTGCTGAAGAAATTCACACTCTCA 2301  
Db 2278 CAACTCTCTGAATTTGGTCAAGTTTTGTAGAACAACTTTGCTGAAGAAATTCACACTCTCA 2337  
Qy 2302 GAATCAGAAATCCCAGGATGTGGAATCTGGGAGGAGAGTGGTCAAGAGCTCTCTGT 2361  
Db 2338 GAATCAGAAATCCCAGGATGTGGAATCTGGGAGGAGAGTGGTCAAGAGCTCTCTGT 2397  
Qy 2362 GGAAGAACAGATAAGAGAAATCGGTATTTATGATGAGGATGAGGATGAGAGTGAACAAAT 2421

Db 2398 GGAAGAACAGATAAGAGAAATCGGTATTTATGATGAGCATGAGGATGAAGAGTGACAAAT 2457  
Qy 2422 TCCAATGATGCTGGGCCCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAA 2481  
Db 2458 TCCAATGATGCTGGGCCCTTAAATTCATGTTAGTGTAGCGAGCCACTGCCCTTTGTCAAA 2517  
Qy 2482 ATGTGATGACATAGCAGGATATCCAGCATGAAATTTACTTGGAAAGTAACTTTG 2541  
Db 2518 ATGTGATGACATAGCAGGATATCCAGCATGAAATTTACTTGGAAAGTAACTTTG 2577  
Qy 2542 GAAAGAAATTCCTTCTTAAATCAAAAAACAAAAACAAAAACACA 2586  
Db 2578 GAAAGAAATTCCTTCTTAAATCAAAAAACAAAAACAAAAACACA 2622

RESULT 3

US-10-037-270-45  
; Sequence 45, Application US/10037270  
; Publication No. US20030104529A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Felyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Tillinghast, John  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030104529A1e1 Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/10/037,270  
; CURRENT FILING DATE: 2002-01-04  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pt\_Fl\_genes Version 1.0  
; SEQ ID NO 45  
; LENGTH: 2905  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (173)..(2452)  
US-10-037-270-45

Query Match 69.2%; Score 2565; DB 15; Length 2905;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2565; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 22 TAGCAGCTTGGTCCGACAGGTGCGCTAGGTAGAGCGCCGGGACCTGTGACAGGGCTGGT 81  
Db 58 TAGCAGCTTGGTCCGACAGGTGCGCTAGGTAGAGCGCCGGGACCTGTGACAGGGCTGGT 117  
Qy 82 AGCAGCGCAGAGGAAAGCGCGCTTTTAGCCAGGATTTTTCAGTGTCTGTAGACAAGATGGA 141  
Db 118 AGCAGCGCAGAGGAAAGCGCGCTTTTAGCCAGGATTTTTCAGTGTCTGTAGACAAGATGGA 177  
Qy 142 ATCATCTCCATTTAATAGACGGCAATGGACCTCACTATCATTTAGGGGTAAACAGCCAAAGA 201  
Db 178 ATCATCTCCATTTAATAGACGGCAATGGACCTCACTATCATTTAGGGGTAAACAGCCAAAGA 237



QY	202	ACCTTCTCTGTCAACAAGAACAAAGTCATCGGCTATTGTGGAAATATTCTCCAGTACCA	261
Db	238	ACTTTTCTCTGTCAACAAGAACAAAGTCATCGGCTATTGTGGAAATATTCTCCAGTACCA	297
QY	262	GAAAGCAGCTGAAGAAACAAACATGAGAGAAAGAGAAAGTAAACCCGAAATCTCTCCCA	321
Db	298	GAAAGCAGCTGAAGAAACAAACATGAGAGAAAGAGAAAGTAAACCCGAAATCTCTCCCA	357
QY	322	GCACCTTTAGAAAGGGGACCCCTGACTGTGTGTTAAAGAAAGTGGGAGAACCCAGGCTGGG	381
Db	358	GCACCTTTAGAAAGGGGACCCCTGACTGTGTGTTAAAGAAAGTGGGAGAACCCAGGCTGGG	417
QY	382	AGCAGAGTCTCACACAGACTCTCTACGGNAACAGCAGCACTGAGATTAGCAGAGCAGA	441
Db	418	AGCAGAGTCTCACACAGACTCTCTACGGNAACAGCAGCACTGAGATTAGCAGAGCAGA	477
QY	442	CCATCCCTGCTGTAAGTGCAGAGCCACCGCTGCTCTGAGGCCAAAGCTGACCAAGAAAGA	501
Db	478	CCATCCCTGCTGTAAGTGCAGAGCCACCGCTGCTCTGAGGCCAAAGCTGACCAAGAAAGA	537
QY	502	ACAAATCCACCCACAGACTAGACTCAGGTCACCTCCTGAAAGCCCTGTTTCAGGGTCGATA	561
Db	538	ACAAATCCACCCACAGACTAGACTCAGGTCACCTCCTGAAAGCCCTGTTTCAGGGTCGATA	597
QY	562	TCCCCACATCAAGGACGGTGAGGATCTTAAAGACACACTCAAAGAAAGTAAAGAAATGGA	621
Db	598	TCCCCACATCAAGGACGGTGAGGATCTTAAAGACACACTCAAAGAAAGTAAAGAAATGGA	657
QY	622	AAATTGCTAGGAGATCCAGGCATCAAGTAGAATAATCAGAAATCAGTGAAACACAGA	681
Db	658	AAATTGCTAGGAGATCCAGGCATCAAGTAGAATAATCAGAAATCAGTGAAACACAGA	717
QY	682	TGCTTCGGGCAAAATAGAGAAATATAATGTTCCGCTGAACAGGCTTAAAGATGATGTTGA	741
Db	718	TGCTTCGGGCAAAATAGAGAAATATAATGTTCCGCTGAACAGGCTTAAAGATGATGTTGA	777
QY	742	GAAAGGTGAACCAACTCAAACCTAAGATTTCTCGGGGCCAAAGCCGAAAGTGAAGTGAAG	801
Db	778	GAAAGGTGAACCAACTCAAACCTAAGATTTCTCGGGGCCAAAGCCGAAAGTGAAGTGAAG	837
QY	802	GAAGATCTCTGAAAAACAGCTATTCTCTAGATGACTGGAAATAGGCCCCAGGTCAGTTGTC	861
Db	838	GAAGATCTCTGAAAAACAGCTATTCTCTAGATGACTGGAAATAGGCCCCAGGTCAGTTGTC	897
QY	862	ATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAATCTCCAGCCCT	921
Db	898	ATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAATCTCCAGCCCT	957
QY	922	CTCAGAAACCTCTATAAAGGATCGAATGGCCAAAGTACCAGGCAGCTGTGTGCCAAACAAAG	981
Db	958	CTCAGAAACCTCTATAAAGGATCGAATGGCCAAAGTACCAGGCAGCTGTGTGCCAAACAAAG	1017
QY	982	CAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGTGGCGAAATCAAAATTCATAA	1041
Db	1018	CAGCTCAACCAACTATACAAATGAGCTGAAAGCCAGTGTGGCGAAATCAAAATTCATAA	1077
QY	1042	AATGGAGCAAAAGGAGAAATCTGCCCCAGGCTCTGAGGCTGCGATCACCCATCAGGAAGS	1101
Db	1078	AATGGAGCAAAAGGAGAAATCTGCCCCAGGCTCTGAGGCTGCGATCACCCATCAGGAAGS	1137
QY	1102	GGAAAGATTTCTGCAAAATCAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCGCAAGATGA	1161
Db	1138	GGAAAGATTTCTGCAAAATCAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCGCAAGATGA	1197
QY	1162	CTCCCGTACTCCAGGTTAAGAGTGAAGTTCACAGCCCTGTCCATCCCAAGCCACTAAG	1221
Db	1198	CTCCCGTACTCCAGGTTAAGAGTGAAGTTCACAGCCCTGTCCATCCCAAGCCACTAAG	1257
QY	1222	TCAGATTCAGAGCCCTCCAGTCTTTCTGAAGTTCTCTCCCAAGCAATGAAGAGTT	1281
Db	1258	TCAGATTCAGAGCCCTCCAGTCTTTCTGAAGTTCTCTCCCAAGCAATGAAGAGTT	1317
QY	1282	TCAGGCACCTGCAAGACAGACCTCGGCTGGAATGTGAGAGACAGTCTATCCAATGGAGCG	1341

Db	1318	TCAGGCACCTGCAAGACAGACCTCGGCTGGAATGTGAGAGACAGTCTATCCAATGGAGCG	1377
QY	1342	TCCTTTGGCCAAACAGCAGGTGTTTACATCAGTGCCTTCGGTTCCTCTATTGCAACAA	1401
Db	1378	TCCTTTGGCCAAACAGCAGGTGTTTACATCAGTGCCTTCGGTTCCTCTATTGCAACAA	1437
QY	1402	CAAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGAAATCTATTGTAAGCCTCACTT	1461
Db	1438	CAAACTCAGTCTAGGAACATATGCATCTTTACATGGAAGAAATCTATTGTAAGCCTCACTT	1497
QY	1462	CAATCAACTCTTTAAATCTAAGGCCAACTATGATGAAGGCTTTTGGGCACAGACCAACAA	1521
Db	1498	CAATCAACTCTTTAAATCTAAGGCCAACTATGATGAAGGCTTTTGGGCACAGACCAACAA	1557
QY	1522	GGATCTATGGCAAGCAAAAAATGAAAAACGAAGAGATTTTGGAGAGACCAAGCCTTGC	1581
Db	1558	GGATCTATGGCAAGCAAAAAATGAAAAACGAAGAGATTTTGGAGAGACCAAGCCTTGC	1617
QY	1582	AAATGCAAGGGAGACCCCTCACAGCCAGGGGTAGAGATGCCCTATTGCTAAGGTGGG	1641
Db	1618	AAATGCAAGGGAGACCCCTCACAGCCAGGGGTAGAGATGCCCTATTGCTAAGGTGGG	1677
QY	1642	TGCTCTGCTCAAGTATGGAAGCCAAAGGCTCTCTCAGCAGGAGAAAGAAACAAAGCC	1701
Db	1678	TGCTCTGCTCAAGTATGGAAGCCAAAGGCTCTCTCAGCAGGAGAAAGAAACAAAGCC	1737
QY	1702	AGCTGAAACCAAGAAAGCTGAGGATCGCTTGCCACCCCCCACTGAACTTGGAAAGTTCAGG	1761
Db	1738	AGCTGAAACCAAGAAAGCTGAGGATCGCTTGCCACCCCCCACTGAACTTGGAAAGTTCAGG	1797
QY	1762	AAGTGCCTTGAGGAAAGGATCAAAATGTCAAAGCCCAATGGCTCTCTGAAAGACGAAAT	1821
Db	1798	AAGTGCCTTGAGGAAAGGATCAAAATGTCAAAGCCCAATGGCTCTCTGAAAGACGAAAT	1857
QY	1822	CAGCAAGCCCAAGTCTCTGAGGATGCGATCTAGATCTGAGAGGCTTAAGACGATCTTC	1881
Db	1858	CAGCAAGCCCAAGTCTCTGAGGATGCGATCTAGATCTGAGAGGCTTAAGACGATCTTC	1917
QY	1882	TTCCTGAGGAAAGAAAGCCGCCCATTCACCTGTAGCAGCTTCATTTCAAAGCACTCTGT	1941
Db	1918	TTCCTGAGGAAAGAAAGCCGCCCATTCACCTGTAGCAGCTTCATTTCAAAGCACTCTGT	1977
QY	1942	CAAGAGCCCAAAAACTGTGTCTCCCACTTATCAGGAAAGGCTGGAGCATGTCCAGAGCAG	2001
Db	1978	CAAGAGCCCAAAAACTGTGTCTCCCACTTATCAGGAAAGGCTGGAGCATGTCCAGAGCAG	2037
QY	2002	TGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAAACAAAGTGGAAATGCCAAGGCTTC	2061
Db	2038	TGAAGAGTCTGTGGGTGGAAGAGTTGCAGAAAGGAAACAAAGTGGAAATGCCAAGGCTTC	2097
QY	2062	TGAAGAGATGGGATGTGGGAAACAAACCTGCAACAAAGAAATCTAAAGGAGAGAC	2121
Db	2098	TGAAGAGATGGGATGTGGGAAACAAACCTGCAACAAAGAAATCTAAAGGAGAGAC	2157
QY	2122	AGGGAAGAGAAAGTAAAGAAAGTTCATAGTTTGGAGATGGAGAAATGAGAAATCTTTGAGAAA	2181
Db	2158	AGGGAAGAGAAAGTAAAGAAAGTTCATAGTTTGGAGATGGAGAAATGAGAAATCTTTGAGAAA	2217
QY	2182	TGGTGCAGACTCCGATGAAGATGATAACAGCTTCTCAAACAAACAAATCTCCACAGAAC	2241
Db	2218	TGGTGCAGACTCCGATGAAGATGATAACAGCTTCTCAAACAAACAAATCTCCACAGAAC	2277
QY	2242	CAAGTCTCTGAATTTGGTTCAGTTTGTAGACAAACCTTTTGTGAGAAATTCACCTACTCA	2301
Db	2278	CAAGTCTCTGAATTTGGTTCAGTTTGTAGACAAACCTTTTGTGAGAAATTCACCTACTCA	2337
QY	2302	GAAATCAGAAATCCAGGATGTGGAGCTCTGGAGGGAGAGTGGTCAAGAGCTCTCTGT	2361
Db	2338	GAAATCAGAAATCCAGGATGTGGAGCTCTGGAGGGAGAGTGGTCAAGAGCTCTCTGT	2397
QY	2362	GGAAGAAACAGATTAAGAGAAATTCGGTATTATGATGAGGATGAGGATGAAGAGTCAACAAAT	2421

Db 2398 GGAAGAACAGATAAAGACAGAAATCGGTATTATGATGAGGATGAGGATGAAGAGTGACAAAT 2457  
QY 2422 TCGCAATGATGCTGGGCTTTAAATTCATGTTACTGTTAGCGCCACTGCCCTTTGTCAAA 2481  
Db 2458 TCGCAATGATGCTGGGCTTTAAATTCATGTTACTGTTAGCGCCACTGCCCTTTGTCAAA 2517  
QY 2482 ATGTGATGCACATAGCAGGATCCAGCATGAAATGTAATTTACTTGGAGTAACCTTTG 2541  
Db 2518 ATGTGATGCACATAGCAGGATCCAGCATGAAATGTAATTTACTTGGAGTAACCTTTG 2577  
QY 2542 GAAAGAAATTCCTTCTTAAATCAAAACCAAAACCAAAACCAAAACCA 2586  
Db 2578 GAAAGAAATTCCTTCTTAAATCAAAACCAAAACCAAAACCAAAACCA 2622

RESULT 4  
US-09-783-732-3  
; Sequence 3, Application US/09783732  
; Publication No. US20030054417A1  
; GENERAL INFORMATION:  
; APPLICANT: Chang, David D.  
; APPLICANT: Maul, Raymond S.  
; APPLICANT: The Regents of the University of California  
; TITLE OF INVENTION: EPITHELIAL PROTEIN LOST IN NEOPLASM  
; FILE REFERENCE: 10809/003001  
; CURRENT APPLICATION NUMBER: US/09/783,732  
; CURRENT FILING DATE: 2001-02-13  
; PRIOR APPLICATION NUMBER: 09/658,400  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 3650  
; TYPE: DNA  
; ORGANISM: Homosapien  
US-09-783-732-3

Query Match 57.5%; Score 2130; DB 11; Length 3650;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3580; Conservative 0; Mismatches 11; Indels 11; Gaps 4;

QY 92 AGGAAAGCGGCTTTTAGCCAGGTATTTCAGTGTCTGTACAGAGATGGAATCATCTCCA 151  
Db 57 AGGAAAGCGGCTTTTAGCCAGGTATTTCAGTGTCTGTACAGAGATGGAATCATCTCCA 116  
QY 152 TTTAATAGACGGCAATGGACCTCACTATCATTGAGGTTACAGCCAAAGAACTTTCTCT 211  
Db 117 TTTAATAGACGGCAATGGACCTCACTATCATTGAGGTTACAGCCAAAGAACTTTCTCT 176  
QY 212 GTCAACAGAAACAGTCAATCGGCTATTGTGAAATATTCTCCAGTACCAGAAAGCAGCT 271  
Db 177 GTCAACAGAAACAGTCAATCGGCTATTGTGAAATATTCTCCAGTACCAGAAAGCAGCT 236  
QY 272 GAAGAAACAAACATGGAGAAAGAGAGTAACACCGAAATCTCTCCAGCACTTTAGA 331  
Db 237 GAAGAAACAAACATGGAGAAAGAGAGTAACACCGAAATCTCTCCAGCACTTTAGA 296  
QY 332 AAGGGACCCCTGACTGTGTTTAAAGAAAGTGGAGAACCCAGGCTGGAGGAGAGTCT 391  
Db 297 AAGGGACCCCTGACTGTGTTTAAAGAAAGTGGAGAACCCAGGCTGGAGGAGAGTCT 356  
QY 392 CACACAGACTCTCTACGGAACAGCAGCACTGAGATTAGGCACAGCAGACCACTCTCT 451  
Db 357 CACACAGACTCTCTACGGAACAGCAGCACTGAGATTAGGCACAGCAGACCACTCTCT 416  
QY 452 GCTGAAGTGACAGGCACTGCTTCTGGAGCCAAAGTGCACCAAGAGAAACAAATCCAC 511  
Db 417 GCTGAAGTGACAGGCACTGCTTCTGGAGCCAAAGTGCACCAAGAGAAACAAATCCAC 476  
QY 512 CCCAGATCTAGACTCAGGTCACTCTCTGAAAGCCCTCGTTTCAGGTCAGATATCCCAATC 571  
Db 477 CCCAGATCTAGACTCAGGTCACTCTCTGAAAGCCCTCGTTTCAGGTCAGATATCCCAATC 536

QY 572 AAGGACGCTGAGGATCTTAAAGACCCTCAACAGAAAGTAAAAAATGAAAAATTTGTCTA 631  
Db 537 AAGGACGCTGAGGATCTTAAAGACCCTCAACAGAAAGTAAAAAATGAAAAATTTGTCTA 596  
QY 632 GGAGAAATCCAGGATGAAGTAGAAAAATCAGAAATCAGTCAAAACACAGATGCTTCGGGC 691  
Db 597 GGAGAAATCCAGGATGAAGTAGAAAAATCAGAAATCAGTCAAAACACAGATGCTTCGGGC 656  
QY 692 AAAATAGAGAAATATAATGTTCCGCTGAACAGGCTTAAGATGATGTTTGAGAAAGGTGAA 751  
Db 657 AAAATAGAGAAATATAATGTTCCGCTGAACAGGCTTAAGATGATGTTTGAGAAAGGTGAA 716  
QY 752 CCAACTCAAACTAGATTCTCCGGGCCCAAGCCGAAGTCAAGTGAAGGAAGATCTCT 811  
Db 717 CCAACTCAAACTAGATTCTCCGGGCCCAAGCCGAAGTCAAGTGAAGGAAGATCTCT 776  
QY 812 GAAACAGCTATTCTCTAGATGACTGGAATAGCCCCAGGTCAAGTCTCATCTTCTACA 871  
Db 777 GAAACAGCTATTCTCTAGATGACTGGAATAGCCCCAGGTCAAGTCTCATCTTCTACA 836  
QY 872 TTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAATCTCCAGGCTCTCAGAAACC 931  
Db 837 TTTGACTCGGAGAAAAATGAGAGTAGACGAAATCTGGAATCTCCAGGCTCTCAGAAACC 896  
QY 932 TCTATAAAGGATCGAATGGCCCAAGTACAGGCAGCTGTGTCCAAACAAAGCAGCTCAACC 991  
Db 897 TCTATAAAGGATCGAATGGCCCAAGTACAGGCAGCTGTGTCCAAACAAAGCAGCTCAACC 956  
QY 992 AACTATACAAATGAGCTGAAAGCCAGTGGTGGCGAAATCAAAATTCATAAATGAGGCAA 1051  
Db 957 AACTATACAAATGAGCTGAAAGCCAGTGGTGGCGAAATCAAAATTCATAAATGAGGCAA 1016  
QY 1052 AAGGAGAAATGTCCCCCAGGTCCTGAGGTCTGCATCAACCCATCAGGAAGGGGAAAGATT 1111  
Db 1017 AAGGAGAAATGTCCCCCAGGTCCTGAGGTCTGCATCAACCCATCAGGAAGGGGAAAGATT 1076  
QY 1112 TCTGCAAAATGAGATAGCTGGCAGTCCGTTCCACCCCTGCGGAAGATGACTCC- --GT 1168  
Db 1077 TCTGCAAAATGAGATAGCTGGCAGTCCGTTCCACCCCTGCGGAAGATGACTCCCAAGT 1136  
QY 1169 GACTCCCAGGTTAAGAGTGAGGTTCAACAGCAGCTGTCCATCCCAAGCCACTAAGTCCAGAT 1228  
Db 1137 GACTCCCAGGTTAAGAGTGAGGTTCAACAGCAGCTGTCCATCCCAAGCCACTAAGTCCAGAT 1196  
QY 1229 TCCAGAGCTCCAGTCTTTCTGAAAGTTCTCTCTCCAAAGCAATCAAGAAAGTTTCAGGCA 1288  
Db 1197 TCCAGAGCTCCAGTCTTTCTGAAAGTTCTCTCTCCAAAGCAATCAAGAAAGTTTCAGGCA 1256  
QY 1289 CTTGCAAGAGAGACCTGCGTGGAAATGTCAAGAGCAGTCTATCCAAATGGAGGCTCTCTTG 1348  
Db 1257 CTTGCAAGAGAGACCTGCGTGGAAATGTCAAGAGCAGTCTATCCAAATGGAGGCTCTCTTG 1316  
QY 1349 GCCAACAGCAGGTTTTCAGATCAGCTGCTTCGTTGCTCTTATTGCAACAAACAACTC 1408  
Db 1317 GCCAACAGCAGGTTTTCAGATCAGCTGCTTCGTTGCTCTTATTGCAACAAACAACTC 1376  
QY 1409 AGTCTAGGAACATATGCACTTTTACATGGAAGAAATCTATTGTAAGCCTCACTTCAATCAA 1468  
Db 1377 AGTCTAGGAACATATGCACTTTTACATGGAAGAAATCTATTGTAAGCCTCACTTCAATCAA 1436  
QY 1469 CTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACCAACAAAGATCTA 1528  
Db 1437 CTCTTTAAATCTAAGGGCAACTATGATGAAGGCTTTGGGCACAGACCAACAAAGATCTA 1496  
QY 1529 TGGGCAAGCAAAATGAAACCGAAGAGATTTTGGAGAGACCCAGCCAGCTTGCATAATGCA 1588  
Db 1497 TGGGCAAGCAAAATGAAACCGAAGAGATTTTGGAGAGACCCAGCCAGCTTGCATAATGCA 1556  
QY 1589 AGGGAGACCCCTCACAGCCAGGGGTAGAGATGCCCTATTGCTAAGGTGGGTGTCCTG 1648  
Db 1557 AGGGAGACCCCTCACAGCCAGGGGTAGAGATGCCCTATTGCTAAGGTGGGTGTCCTG 1616

1649 GCTGCAAGTATGGAAGCAAGGCTCTCTCAGCAGGAGAAAGGAGACAGCCAGCTGAA 1708  
1617 GCTGCAAGTATGGAAGCAGGAGGCTCTCTCAGCAGGAGAAAGGAGACAGCCAGCTGAA 1676  
1709 ACCAAGAGCTGAGGATGCGCTGGCCACCCGCCACTGAACTTGGAAAGTTTCAGGAAGTGC 1768  
1677 ACCAAGAGCTGAGGATGCGCTGGCCACCCGCCACTGAACTTGGAAAGTTTCAGGAAGTGC 1736  
1769 TTGGAGGAAGGATCAAAATGTCAAGGCCAAATGGCCCTCTGAGAGCAAAATCAGCAAG 1828  
1737 TTGGAGGAAGGATCAAAATGTCAAGGCCAAATGGCCCTCTGAGAGCAAAATCAGCAAG 1796  
1829 CCCGAAGTCTGAGGATGTCGATCTAGATCTGGAAGAGCTAAGACGATCTTCTTCACTG 1888  
1797 CCCGAAGTCTGAGGATGTCGATCTAGATCTGGAAGAGCTAAGACGATCTTCTTCACTG 1856  
1889 AAGGAAGAGCGGCCCATCTCACTGTAGCAGCTTCAATTTCAAAGCACCTCTGTCAAAGAGC 1948  
1857 AAGGAAGAGCGGCCCATCTCACTGTAGCAGCTTCAATTTCAAAGCACCTCTGTCAAAGAGC 1916  
1949 CCAAAACCTGTGTCCCACTATCAGGAAGGCTGGAGCATGTGAGAGCAAAATCAGCAAG 2008  
1917 CCAAAACCTGTGTCCCACTATCAGGAAGGCTGGAGCATGTGAGAGCAAAATCAGCAAG 1976  
2009 TCTGTGGTGGAGAGTTGCGAAGAGAAACAGTGGGAAATGCCAAGGCTTCTTAAGAG 2068  
1977 TCTGTGGTGGAGAGTTGCGAAGAGAAACAGTGGGAAATGCCAAGGCTTCTTAAGAG 2036  
2069 AATGGGAATCTGGGAAACCAACCTGGCAAAACAAAGAAATCTAAAGGAGAGACAGGGAAG 2128  
2037 AATGGGAATCTGGGAAACCAACCTGGCAAAACAAAGAAATCTAAAGGAGAGACAGGGAAG 2096  
2129 AGAAGTAAGGAAGTGTATAGTTGGAGATGGAGAAATGGAATCTTTGAGAAATGGTGCA 2188  
2097 AGAAGTAAGGAAGTGTATAGTTGGAGATGGAGAAATGGAATCTTTGAGAAATGGTGCA 2156  
2189 GACTCCGATGAAAGATGATAACAGCTTCTCAACACAACTCCCAAGAACCCAGTCT 2248  
2157 GACTCCGATGAAAGATGATAACAGCTTCTCAACACAACTCCCAAGAACCCAGTCT 2216  
2249 CTGAATGTGTCGAGTTTGTAGACACACCTTGTGAGAAATCTACTACTCAGAATCAG 2308  
2217 TTGAATTTGTCGAGTTTGTAGACACACCTTGTGAGAAATCTACTACTCAGAATCAG 2276  
2309 AATCCCGAGATGGAACCTCTGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAA 2368  
2277 AATCCCGAGATGGAACCTCTGGAGGAGAGTGGTCAAGAGCTCTCTGTGGAAGAA 2336  
2369 CAGATAAAGAGAAATCGGTATTTATGATGAGATGAGATGAAAGTGACAAATTTGCAATG 2428  
2337 CAGATAAAGAGAAATCGGTATTTATGATGAGATGAGATGAAAGTGACAAATTTGCAATG 2396  
2429 ATGCTGGCCCTTAAATTCATGTTAGTGTAGCAGCCACTGCCCTTTGTCAAAATGTGAT 2488  
2397 ATGCTGGCCCTTAAATTCATGTTAGTGTAGCAGCCACTGCCCTTTGTCAAAATGTGAT 2456  
2489 GCACATAGCAGGATCCACGATGAAATGTAATTTTACTTTGGAAGTAACTTTGGAAGAA 2548  
2457 GCACATAGCAGGATCCACGATGAAATGTAATTTTACTTTGGAAGTAACTTTGGAAGAA 2516  
2549 ATTCCCTCTTAAATTCAAATAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2608  
2517 ATTCCCTCTTAAATTCAAATAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 2576  
2609 AGATAACCTTTACTTAAATTTCTTATTTAGAGTGTATGATATGATAGTCTGTAAAGGC 2668  
2577 AGATAACCTTTACTTAAATTTCTTATTTAGAGTGTATGATATGATAGTCTGTAAAGGC 2631  
2669 TTGTAACCTGGGAAATATTCACCTGATAAATAGCCCAATCTACTGTATTTCCCAAAAGG 2728  
2632 TTGTAACCTGGGAAATATTCACCTGATAAATAGCCCAATCTACTGTATTTCCCAAAAGG 2691  
2729 CAATATTAAAGGTAGATGATATTAGTAGTATATTGTTACACACTATTTTGGAAATTAGAG 2788

2692 CAATATTAAAGGTAGATGATGATTAGTAGTATATTGTTACACACTATTTTGGAAATTAGAG 2751  
2789 AACATACAGAAGAAATTTAGGGGCTTTAAACATTTACGACTGAATGCACTTTAGTATAAAGG 2848  
2752 AACATACAGAAGAAATTTAGGGGCTTTAAACATTTACGACTGAATGCACTTTAGTATAAAGG 2810  
2849 GCACAGTTGTATATTTTAAATGAATACCAATTTAAATTTTAAATTTTAAATTTTAAATTT 2908  
2811 GCACAGTTGTATATTTTAAATGAATACCAATTTAAATTTTAAATTTTAAATTTTAAATTT 2870  
2909 AGATTATTTAGTCTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2968  
2871 AGATTATTTAGTCTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 2930  
2969 TTACTACTTTATGTCGTCTCTCTAAACTACATCTCGAATCGACGTCCTGAGGATATAA 3028  
2931 TTACTACTTTATGTCGTCTCTCTAAACTACATCTCGAATCGACGTCCTGAGGATATAA 2988  
3029 ACAACAGAGCACCTTTTGGAGCAATTTGAGGCAATTTGAGGCAATTTGAGGCAATTTGAG 3088  
2989 ACAACAGAGCACCTTTTGGAGCAATTTGAGGCAATTTGAGGCAATTTGAGGCAATTTGAG 3048  
3089 GATCTGTCTCTCCCAATTAAGCTTTTGTATCTGCCAGTGAATTTACTGACTCCAAATG 3148  
3049 GATCTGTCTCTCCCAATTAAGCTTTTGTATCTGCCAGTGAATTTACTGACTCCAAATG 3108  
3149 ATTGCTTTCTTTCTGTCGTATCTGTCCTCTCATTAATTTACTGAAAGCTGCAATATTTT 3208  
3109 ATTGCTTTCTTTCTGTCGTATCTGTCCTCTCATTAATTTACTGAAAGCTGCAATATTTT 3168  
3209 AGTAATACCTTTGGGATCACCTGTCCTCCCATCTTCCGTTTGTAGACAAAGTGAAGAGTTT 3268  
3169 AGTAATACCTTTGGGATCACCTGTCCTCCCATCTTCCGTTTGTAGACAAAGTGAAGAGTTT 3228  
3269 AAGGAGGAAGAAAGAACTGTTTACACCACTTTAGCTCAGACCTCTAAACCTGTAT 3328  
3229 AAGGAGGAAGAAAGAACTGTTTACACCACTTTAGCTCAGACCTCTAAACCTGTAT 3288  
3329 TTCCCTTATGATGTCCCTTTTGGAGCACTAAATTTTAAATTTTAAATTTTAAATTTTAA 3388  
3289 TTCCCTTATGATGTCCCTTTTGGAGCACTAAATTTTAAATTTTAAATTTTAAATTTTAA 3348  
3389 ATATTGATTTTATCACAGTATCTCAGGTTGAAATTTAAACCACTATAGGCTTTTCT 3448  
3349 ATATTGATTTTATCACAGTATCTCAGGTTGAAATTTAAACCACTATAGGCTTTTCT 3408  
3449 TGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACCTTGAGTACATTTGTTGTA 3508  
3409 TGGGATGATTTTCTAGTCTTAAGGTTTGGGACATTTAAACCTTGAGTACATTTGTTGTA 3468  
3509 CACAGTTGATATTCACAAATTTGATGATGGGAGGAGGTTCTTTAAGCTGTAGGCTTT 3568  
3469 CACAGTTGATATTCACAAATTTGATGATGGGAGGAGGTTCTTTAAGCTGTAGGCTTT 3528  
3569 TCTTTGTACTGCAATTTATAGAGATTTAGCTTTAAATTTTAAATTTTAAATTTTAAATTT 3628  
3529 TCTTTGTACTGCAATTTATAGAGATTTAGCTTTAAATTTTAAATTTTAAATTTTAAATTT 3588  
3629 GCTTTCTTACTGCTTACCTAGTCTGAAACATTTTAAATTTTAAATTTTAAATTTTAAATTT 3688  
3589 GCTTTCTTACTGCTTACCTAGTCTGAAACATTTTAAATTTTAAATTTTAAATTTTAAATTT 3648  
3689 TG 3690  
3649 TG 3650

RESULT 5  
US-09-783-732-1  
; Sequence 1, Application US/09783732  
; Publication No. US20030054417A1  
; GENERAL INFORMATION:

APPLICANT: Chang, David D.  
APPLICANT: Maul, Raymond S.  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: EPITHELIAL PROTEIN LOST IN NEOPLASM  
FILE OF INVENTION: (BEPIN)  
FILE REFERENCE: 10809/003001  
CURRENT APPLICATION NUMBER: US/09/783,732  
CURRENT FILING DATE: 2001-02-13  
PRIOR APPLICATION NUMBER: 09/658,400  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 3543  
TYPE: DNA  
ORGANISM: Homosapien  
us-09-783-732-1

Query Match 55.1%; Score 2041; DB 11; Length 3543;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3371; Conservative 11; Indels 10; Gaps 3;  
302 AACACCGGAAATCTCTCCAGCAGCTTTAGAAAGGGGACCCCTGACTGTGTTAAAGAGAAG 361  
159 AACACCGGAAATCTCTCCAGCAGCTTTAGAAAGGGGACCCCTGACTGTGTTAAAGAGAAG 218  
362 TGGGAGAACCCAGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGAACAGCAGCACT 421  
219 TGGGAGAACCCAGGCTGGGAGCAGAGTCTCACACAGACTCTCTACGGAACAGCAGCACT 278  
422 GAGATTAGGCACAGACAGACATCTCTCTGCTGAAAGTGAACGCCACGCTGCTTCTGGA 481  
279 GAGATTAGGCACAGACAGACATCTCTCTGCTGAAAGTGAACGCCACGCTGCTTCTGGA 338  
482 GCCAAGGTGACCAAGAGAACAAATCACACCCAGATCTAGACTCAGGTCACTCTCTGAA 541  
339 GCCAAGGTGACCAAGAGAACAAATCACACCCAGATCTAGACTCAGGTCACTCTCTGAA 398  
542 GCCCTCGTTCAGGTCGATATCCACACATCAAGACGCTGAGGATCTTAAAGACCACTCA 601  
399 GCCCTCGTTCAGGTCGATATCCACACATCAAGACGCTGAGGATCTTAAAGACCACTCA 458  
602 ACAGAAAGTAAAAAATGAAAAATTTCTAGAGAAATCCAGGCATGAAGTGAAGAAATCA 661  
459 ACAGAAAGTAAAAAATGAAAAATTTCTAGAGAAATCCAGGCATGAAGTGAAGAAATCA 518  
662 GAATCAGTGAAGAACACAGATCTTCGGGCAAAATAGAGAAATATAATGTTCCGCTGAAC 721  
519 GAATCAGTGAAGAACACAGATCTTCGGGCAAAATAGAGAAATATAATGTTCCGCTGAAC 578  
722 AGCCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAGATCTCCGGGCCCAA 781  
579 AGCCTTAAGATGATGTTTGAAGAAAGTGAACCACTCAAACTAAGATCTCCGGGCCCAA 638  
782 AGCAGAGTGAAGTGAAGAGGATCTCTGAAACAGACTATTTCTAGATCACTCTGAA 841  
639 AGCAGAGTGAAGTGAAGAGGATCTCTGAAACAGACTATTTCTAGATCACTCTGAA 698  
842 ATAGGCCCCAGGTGAGTGTGATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACA 901  
699 ATAGGCCCCAGGTGAGTGTGATCTTCTACATTTGACTCGGAGAAAAATGAGAGTAGACA 758  
902 AATCTGGAATCTCCAGCCTCTCAGAAACCTCTATAAGAGTGAATGCGCAAGTACCAG 961  
759 AATCTGGAATCTCCAGCCTCTCAGAAACCTCTATAAGAGTGAATGCGCAAGTACCAG 818  
962 GCAGCTGTGTCACAAAGAGAGCTCAACCACTATACAAATGAGTGAAGCCAGTGGT 1021  
819 GCAGCTGTGTCACAAAGAGAGCTCAACCACTATACAAATGAGTGAAGCCAGTGGT 878  
1022 GGCAGAAATCAAAATTCATAAAATGGAGCAAAAGAGAAATGTGCCCCAGGTCCTGAGTGC 1081  
879 GGCAGAAATCAAAATTCATAAAATGGAGCAAAAGAGAAATGTGCCCCAGGTCCTGAGTGC 938

1082 TGCATACCCCATCAGGAAGGGGAAAAAGATTTCTCAAATGAGAAATAGCCTGGCAGTCCGT 1141  
939 TGCAATACCCCATCAGGAAGGGGAAAAAGATTTCTCAAATGAGAAATAGCCTGGCAGTCCGT 998  
1142 TCCACCCCTGCGGAAGATGACTCCC---GTCACTCCCAGGTTAAGAGTGAGGTTCAACAG 1198  
999 TCCACCCCTGCGGAAGATGACTCCCAGGTTAAGAGTGAGGTTCAACAG 1058  
1199 CCGTCCATCCCAAGCCCACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTCTGAAAGTTCT 1258  
1059 CCGTCCATCCCAAGCCCACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTCTGAAAGTTCT 1118  
1259 CCGTCCATCCCAAGCCCACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTCTGAAAGTTCT 1318  
1119 CCGTCCATCCCAAGCCCACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTCTGAAAGTTCT 1178  
1319 AAGACAGTCTATCCCAATGGAGCGTCTCTTGGCCAAACACAGAGAGCTGCTGGAATGTCTAG 1378  
1179 AAGACAGTCTATCCCAATGGAGCGTCTCTTGGCCAAACACAGAGAGCTGCTTTCACATCAGTGC 1238  
1379 TCCGTTGCTCTATTGCAACCAAACTCAGTCTAGGAACATATGCAATCTTTACATGGA 1438  
1239 TCCGTTGCTCTATTGCAACCAAACTCAGTCTAGGAACATATGCAATCTTTACATGGA 1298  
1439 AGAATCTATTGAGCCTCACTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGAA 1498  
1299 AGAATCTATTGAGCCTCACTTCAATCAACTCTTTAAATCTAAGGGCAACTATGATGAA 1358  
1499 GGCTTTGGGCACAGACCAACAGGATCTATGGCAAGCAAAATGAAACCAAGAGATT 1558  
1359 GGCTTTGGGCACAGACCAACAGGATCTATGGCAAGCAAAATGAAACCAAGAGATT 1418  
1559 TTGGAGAGACCCAGCCAGCTTGCAAAATGCAAGGGAGACCCCTCACAGCCAGGGGTAGAA 1618  
1419 TTGGAGAGACCCAGCCAGCTTGCAAAATGCAAGGGAGACCCCTCACAGCCAGGGGTAGAA 1478  
1619 GATGCCCTATTCTAAGTGGGTGCTTGGCTGCAAGTATGGAAGCCAGGCTCTCTCT 1678  
1479 AATGCCCTATTCTAAGTGGGTGCTTGGCTGCAAGTATGGAAGCCAGGCTCTCTCT 1538  
1679 CAGCAGGAGAGAGACAGCCAGCTGAAACCAAGAGCTGAGGATCGCTCGCCACCC 1738  
1539 CAGCAGGAGAGAGACAGCCAGCTGAAACCAAGAGCTGAGGATCGCTCGCCACCC 1598  
1739 CCACACTGAACTTGGAAGTTGAGGAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCC 1798  
1599 CCACACTGAACTTGGAAGTTGAGGAAGTCCCTTGGAGGAAGGATCAAAATGTCAAAGCCC 1658  
1799 AAATGGCCTCTCTGAAGACGAAATCAGCAAGCCGAAAGTTCCTGAGGATGTCGATCTAGAT 1858  
1659 AAATGGCCTCTCTGAAGACGAAATCAGCAAGCCGAAAGTTCCTGAGGATGTCGATCTAGAT 1718  
1859 CTGAAGAGCTTAAGACGATCTTCTTCAAGGAAAGAGCCGCCATTCACCTGTAGCA 1918  
1719 CTGAAGAGCTTAAGACGATCTTCTTCAAGGAAAGAGCCGCCATTCACCTGTAGCA 1778  
1919 GCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGGAAA 1978  
1779 GCTTCATTTCAAAGCACCTCTGTCAAGAGCCCAAAACTGTGTCCCACTATCAGGAAA 1838  
1979 GGCTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGTGGAAAGAGTTGAGAAAGAAA 2038  
1839 GGCTGGAGCATGTCAGAGCAGAGTGAAGAGTCTGTGGTGGAAAGAGTTGAGAAAGAAA 1898  
2039 CAATGGAAAAATGCCAAGGCTTTTAAGAAAGTGGGAAATGGGAAAAAACAACCTGGCAA 2098  
1899 CAATGGAAAAATGCCAAGGCTTTTAAGAAAGTGGGAAATGGGAAAAAACAACCTGGCAA 1958  
2099 AACAAGAGTCTAAGAGGAGACAGAGGAGAGTGAAGAGTCTAGTTTGGAGATG 2158  
1959 AACAAGAGTCTAAGAGGAGACAGAGGAGAGTGAAGAGTCTAGTTTGGAGATG 2018

2159 GAGAAATGAGATCTTTAGAAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCCTC 2218  
2019 GAGAAATGAGAAATTTTGTAGAAAATGGTGCAGACTCCGATGAAGATGATAACAGCTTCCCTC 2078  
2219 AAACAACAATCTCCACAGAAACCCAAAGTCTCTGAATTTGGTCGAGTTTGTAGACAACACC 2278  
2079 AAACAACAATTTCCACAGAAACCCAAAGTCTCTGAATTTGGTCGAGTTTGTAGACAACACC 2138  
2279 TTTGCTGAAGAAATTCATCTACTCAGATCAGAAATCCCAAGGATGGGAACTCTGGAGGGA 2338  
2139 TTTGCTGAAGAAATTCATCTACTCAGATCAGAAATCCCAAGGATGGGAACTCTGGAGGGA 2198  
2339 GAAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAGAAATCGGTATTATGATGAG 2398  
2199 GAAGTGGTCAAGAGCTCTCTGTGGAAGAACAGATAAAGAGAGAAATCGGTATTATGATGAG 2258  
2399 GATGAGGATGAAGAGTGAACAAATTCGAATGATGCTGGCCCTTAAATTCATGTTAGTGTGA 2458  
2259 GATGAGGATGAAGAGTGAACAAATTCGAATGATGCTGGCCCTTAAATTCATGTTAGTGTGA 2318  
2459 GCGAGCCACTGCCCCCTTTGTCAAAATGTGATGACATAGCAGATATCCAGCATGAAATG 2518  
2319 GCGAGCCACTGCCCCCTTTGTCAAAATGTGATGACATAGCAGATATCCAGCATGAAATG 2378  
2519 TAATTTACTTGGAAAGTAACTTTGGAAAGAAATTCCTTCTTAAATCAAAAACAAAACAAA 2578  
2379 TAATTTACTTGGAAAGTAACTTTGGAAAGAAATTCCTTCTTAAATCAAAAACAAAACAAA 2438  
2579 AAACAACAACAACAATCTTAATTAATCTAGAGATAAATTTACTTAAATTTCTTCATTTAG 2638  
2439 AAACAACAACAACAATCTTAATTAATCTAGAGATAAATTTACTTAAATTTCTTCATTTAG 2493  
2639 CAGTGATGATATGATCAATGATGCTTAAGGCTTGTAACTGGGAAATATTTCCACCTGATAA 2698  
2494 CAGTGATGATATGATCAATGATGCTTAAGGCTTGTAACTGGGAAATATTTCCACCTGATAA 2553  
2699 TAGCCAGATTTCTACTGTTATTTCCAAAAGCAATATTAAGGTAGATGATGATGATGATG 2758  
2554 TAGCCAGATTTCTACTGTTATTTCCAAAAGCAATATTAAGGTAGATGATGATGATGATG 2613  
2759 ATATTGTACACACTATTTTGGAAATTAAGAAATAGAAATTAAGGCTTTAAAC 2818  
2614 ATATTGTACACACTATTTTGGAAATTAAGAAATAGAAATTAAGGCTTTAAAC 2673  
2819 ATTACGACTGAATCCACTTTAGTATAAAGGCGCAGTTTGTATATTTTAAATGAATACC 2878  
2674 ATTACGACTGAATCCACTTTAGTATAAAGGCGCAGTTTGTATATTTTAAATGAATACC 2733  
2879 AATTTAATTTTGTAGTATTTACCTGTTAAGAGATTTATTTAGTCTTTAAATTTTGTAGTT 2938  
2734 AATTTAATTTTGTAGTATTTACCTGTTAAGAGATTTATTTAGTCTTTAAATTTTGTAGTT 2793  
2939 AATTTTCTGTGTGATATATAGAGAAATTTACTTATTTATGCTCTCTCTAAACTA 2998  
2794 AATTTTCTGTGTGATATATAGAGAAATTTACTTATTTATGCTCTCTCTAAACTA 2853  
2999 CATCTGAACCTGAGCTCTCGAGTATATTAACAACAGAGCACTTTTGGAGCAATGAAA 3058  
2854 CATCTGAACCTGAGCTCTCGAGTATATTAACAACAGAGCACTTTTGGAGCAATGAAA 2911  
3059 AACCACCTACACTCTTCGGTGTCTAGAGAGATCTGCTGTCTCCAAATAAGCTTTTGTGA 3118  
2912 AACCACCTACACTCTTCGGTGTCTAGAGAGATCTGCTGTCTCCAAATAAGCTTTTGTGA 2971  
3119 TCTGCCAGTGAATTTACTGTACTCCAAATGATTTCTTTCTGCGTGTATCTGTGCT 3178  
2972 TCTGCCAGTGAATTTACTGTACTCCAAATGATTTCTTTCTGCGTGTATCTGTGCT 3031  
3179 TCTCATATTTACTGAAGCTGAATATTTTGTATATCTTCGGGATCACTGTCCCCCAT 3238  
3032 TCTCATATTTACTGAAGCTGAATATTTTGTATATCTTCGGGATCACTGTCCCCCAT 3091  
3239 CTTCCGTGTTAGAGCAAAAGTGAAGAGTTTAAAGGAGGAAGAAAGAACTGTCTTACAC 3298

3092 CTTCCGTGTTAGAGCAAGTGAAGAGTTTAAAGGAGGAAGAAAGAACTGTCTTACAC 3151  
3299 CACTTCAGCTCAGACCTCTAAACCCTGTATTTCCCTTATGATGTCCTCTTTTGTAGACAC 3358  
3152 CACTTCAGCTCAGACCTCTAAACCCTGTATTTCCCTTATGATGTCCTCTTTTGTAGACAC 3211  
3359 TAATTTTAAATTTACTTACTAGCTCTGAAATATATTTTATTTATCAGATTTTCTCAGGG 3418  
3212 TAATTTTAAATTTACTTACTAGCTCTGAAATATATTTTATTTATCAGATTTTCTCAGGG 3271  
3419 TGAATTTAAACCACCACTATAGGCTTTTCTTTGGGATGATTTTCTAGTCTTTAAAGTTTGGG 3478  
3272 TGAATTTAAACCACCACTATAGGCTTTTCTTTGGGATGATTTTCTAGTCTTTAAAGTTTGGG 3331  
3479 GACATTATAAACCITGAGTACATTTGTTGTACACAGTTGATATTTCCAAATTTGTATGGATGG 3538  
3332 GACATTATAAACCITGAGTACATTTGTTGTACACAGTTGATATTTCCAAATTTGTATGGATGG 3391  
3539 GAGGAGAGGTGCTTTAAAGCTGTAGGCTTTTCTTTGTACTGCAATTTATAGAGATTTAGCT 3598  
3392 GAGGAGAGGTGCTTTAAAGCTGTAGGCTTTTCTTTGTACTGCAATTTATAGAGATTTAGCT 3451  
3599 TTAATATTTTATAGAGATGTAACACATTTCTGCTTTCTTTAGTCTTACCTAGTCTGAAACAT 3658  
3452 TTAATATTTTATAGAGATGTAACACATTTCTGCTTTCTTTAGTCTTACCTAGTCTGAAACAT 3511  
3659 TTTTATTTCAATAAAGATTTTAAATTTAAATTTG 3690  
3512 TTTTATTTCAATAAAGATTTTAAATTTAAATTTG 3543

RESULT 6  
US-09-925-297-152  
; Sequence 152, Application US/09925297  
; Patent No. US20020081659A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA105  
; CURRENT APPLICATION NUMBER: US/09/925,297  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05989  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 928  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 152  
; LENGTH: 732  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (729)  
; OTHER INFORMATION: n equals a,t,g, or c  
; US-09-925-297-152

Query Match 15.0%; Score 556; DB 9; Length 732;  
Best Local Similarity 99.6%; Pred. No. 1.6e-256;  
Matches 706; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1096 GGAAGGGGAAAGATTTCTGCAATAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCCGA 1155  
Db 1 GGAAGGGGAAAGATTTCTGCAATAGAAATAGCCTGGCAGTCCGTTCCACCCCTGCCGA 60  
QY 1156 AGATGACTCCCGTACTCCCGAGTTAAGAGTGAAGTTCAACAGCCTGTCCATCCCAAGCC 1215  
Db 61 AGATGACTCCCGTACTCCCGAGTTAAGAGTGAAGTTCAACAGCCTGTCCATCCCAAGCC 120  
QY 1216 ACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTTCTGAAAGTTCTCTCCCAAAAGCAATGAA 1275  
Db 121 ACTAAGTCCAGATTTCCAGAGCCTCCAGTCTTTCTGAAAGTTCTCTCTCCCAAAAGCAATGAA 180

	Matches	473;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
1276	GAAGTTTCAGGACCTGCGTGAAGAGACAGACCTGCGTGAAGATGTCAGAGACAGTCTATCCAAT	1335								
181	GAAGTTTCAGGACCTGCGTGAAGAGACAGACCTGCGTGAAGATGTCAGAGACAGTCTATCCAAT	240								
1336	GGAGCGTCTTGGCCAAACAGCAGAGGTGTTTACATCAGCTGCTTCCGTTGCTCTATTG	1395								
241	GGAGCGTCTTGGCCAAACAGCAGAGGTGTTTACATCAGCTGCTTCCGTTGCTCTATTG	300								
1396	CAACAACAACCTAGTCTAGGAACATATGATCTTTACATGGAAGAAATCTATTGTAAGCC	1455								
301	CAACAACAACCTAGTCTAGGAACATATGATCTTTACATGGAAGAAATCTATTGTAAGCC	360								
1456	TCACCTCAATCAACTCTTTAAATCTAAGGCACTATGATCAAGGCTTTTGGGCACAGACC	1515								
361	TCACCTCAATCAACTCTTTAAATCTAAGGCACTATGATCAAGGCTTTTGGGCACAGACC	420								
1516	ACACAAGATCTATGGCAAGCAAAATGAAACGAAAGATTTTGGAGAGACAGGCCA	1575								
421	ACACAAGATCTATGGCAAGCAAAATGAAACGAAAGATTTTGGAGAGACAGGCCA	480								
1576	GCTTCAATCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTATTGCTAA	1635								
481	GCTTCAATCAAGGAGACCCCTCACAGCCAGGGGTAGAAGATGCCCTATTGCTAA	540								
1636	GGTGGGTGCTTGGCAAGCAAAATGAAACGAAAGATTTTGGAGAGACAGGCCA	1695								
541	GGTGGGTGCTTGGCAAGCAAAATGAAACGAAAGATTTTGGAGAGACAGGCCA	600								
1696	CAAGCCAGCTGAAACCAAGCAAGATGAGGATCGCTGCGCCACCCCACTGAACTTGGAG	1755								
601	CAAGCCAGCTGAAACCAAGCAAGATGAGGATCGCTGCGCCACCCCACTGAACTTGGAG	660								
1756	TTCAAGAAAGTGCCTTGGAGGAAGGATCAAAATGTCAGAACCCCAATGG	1804								
661	TTCAAGAAAGTGCCTTGGAGGAAGGATCAAAATGTCAGAACCCCAATGG	709								

RESULT 7  
US-10-066-543-3200  
; Sequence 3200, Application US/10066543  
; Publication No. US20030087818A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yugu  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiaqun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Secrist, Heather  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margrita  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3200  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 533  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-3200

Query Match 11.4%; Score 423; DB 15; Length 547;  
Best Local Similarity 99.8%; Pred. No. 4.5e-192;  
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1925 TTTCAGGACCTCTGTCAAGAGCCCAAAACTGTGTCCACCTATCAGGAAAGGCTGG 1984  
Db 1 TTTCAGGACCTCTGTCAAGAGCCCAAAACTGTGTCCACCTATCAGGAAAGGCTGG 60  
QY 1985 ACCATCTCAGACAGTGAAGAGTCTGTGGTGAAGAGTTCAGAAAGAAACAGTG 2044

Db 215 TACACACTATTTTGGAAATAGAGAAATACAGAGAAATTTAGGGCTTAAACATTACGA 274  
QY 2826 CTGAATGCACCTTTAGTATAAAGGGCACAGTTGTATATTTTAAATGAATACCAATTTAA 2885  
Db 275 CTGAATGCACCTTTAGTATAAAGGGCACAGTTGTATATTTTAAATGAATACCAATTTAA 334  
QY 2886 TTTTATAGTATTTACCTGTTAAGAGATTTAGTCTTTTAAATTTTATAGTTTAAATTTT 2945  
Db 335 TTTTATAGTATTTACCTGTTAAGAGATTTAGTCTTTTAAATTTTATAGTTTAAATTTT 394  
QY 2946 TTGCTGTGATATATAGAGAAATTTAGTCTTTTAAATTTTATAGTTTAAATTTT 3005  
Db 395 TTGCTGTGATATATAGAGAAATTTAGTCTTTTAAATTTTATAGTTTAAATTTT 454  
QY 3006 AACTCGAGCTCTGAGGTATATATACACAGAG 3037  
Db 455 AACTCGAGCTCTGAGGTATATATACACAGAG 486

RESULT 10  
US-10-198-846-10895  
; Sequence 10895, Application US/10198846  
; Publication No. US2003099974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10895  
; LENGTH: 698  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1, 2, 691, 692, 693, 694, 695, 696, 697, 698  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-10895

Query Match 10.4%; Score 385; DB 15; Length 698;  
Best Local Similarity 100.0%; Pred. No. 2.9e-174;  
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 AACACCGAAATCTCTCCAGCACCTTTAGAAAGGGGACCTGACTGTGTTAAAGAGAGAG 361  
Db 293 AACACCGAAATCTCTCCAGCACCTTTAGAAAGGGGACCTGACTGTGTTAAAGAGAGAG 352  
QY 362 TGGGAGAACCCAGGGCTGGAGCAGAGTCTCACAGACTCTCTACGGAACAGCAGCACT 421  
Db 353 TGGGAGAACCCAGGGCTGGAGCAGAGTCTCTACGGAACAGCAGCACT 412  
QY 422 GAGATTAGGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 481  
Db 413 GAGATTAGGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 472  
QY 482 GCCAAGCTGACCAAGAGAAACAAATCCACCCAGATCTTAGCTCAGGTCACTCTCTGAA 541  
Db 473 GCCAAGCTGACCAAGAGAAACAAATCCACCCAGATCTTAGCTCAGGTCACTCTCTGAA 532  
QY 542 GCCCTCCTTCAAGGTCGATATATCCACATCAAGACGGTGAGGATCTTAAGACCACTCA 601  
Db 533 GCCCTCCTTCAAGGTCGATATATCCCAATCAAGACGGTGAGGATCTTAAGACCACTCA 592

b 61 AGCATGTGAGAGAGTGAAGAGTCTGTGGTGGAGAGTTGAGAAAGAAACAGTG 120  
y 2045 GAAATGCAAGCTTCTTAAGAGAGTGAAGTGTGGGAAAAACAACCTTGGCAAAACAAA 2104  
b 121 GAAATGCAAGCTTCTTAAGAGAGTGAAGTGTGGGAAAAACAACCTTGGCAAAACAAA 180  
y 2105 GAATCTAAAGGAGACAGGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 2164  
b 181 GAATCTAAAGGAGAGACAGGAGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 240  
y 2165 GAGAATCTTGTAGAAATGTTGAGAGTCTCGATGAAGATGATACAGCTTCTCTCAACAA 2224  
b 241 GAGAATCTTGTAGAAATGTTGAGAGTCTCGATGAAGATGATACAGCTTCTCTCAACAA 300  
y 2225 CAATCTCCCAAGAACCAAGTCTCAATTTGTTGAGAGTCTCTAGACACACCTTTGCT 2284  
b 301 CAATCTCCCAAGAACCAAGTCTCTCAATTTGTTGAGAGTCTCTAGACACACCTTTGCT 360  
y 2285 GAAGAATCTCACTACTCAGATCAGAAATCCAGGATGGAAGTCTCTGGAGGGAG 2339  
b 361 GAAGAATCTCACTACTCAGATCAGAAATCCAGGATGGAAGTCTCTGGAGGGAG 415

RESULT 9  
US-10-066-543-3373  
; Sequence 3373, Application US/10066543  
; Publication No. US2003008781A1  
; GENERAL INFORMATION:  
; APPLICANT: Jiang, Yuhui  
; APPLICANT: Pyle, Ruth A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Indrias, Carol Yoseph  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Sectist, Heather  
; APPLICANT: Carter, Derrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Smith, Carole L.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Stolk, John A.  
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
; FILE REFERENCE: 21021.563  
; CURRENT APPLICATION NUMBER: US/10/066,543  
; CURRENT FILING DATE: 2002-01-31  
; NUMBER OF SEQ ID NOS: 3417  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3373  
; LENGTH: 583  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 534  
; OTHER INFORMATION: n = A,T,C or G  
US-10-066-543-3373

Query Match 10.8%; Score 401; DB 15; Length 583;  
Best Local Similarity 99.8%; Pred. No. 5.8e-182;  
Matches 451; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2586 AAAAACAACATTTCTAAATCTAGAGTAACCTTACTTAAATCTCTTCTTCTTCTTCTTCT 2645  
Db 35 AAAAACAACATTTCTAAATCTAGAGTAACCTTACTTAAATCTCTTCTTCTTCTTCTTCT 94  
QY 2646 GATATGCATAGTCTGTAAGGCTTGAATCTGAGGAAATATTTCCACCTGATAATAGCCCA 2705  
Db 95 GATATGCATAGTCTGTAAGGCTTGAATCTGAGGAAATATTTCCACCTGATAATAGCCCA 154  
QY 2706 GATCTCTAGTATTTCCCAAGGCAATATTAAGGTAGATAGATAGTATAGTATATTTGT 2765  
Db 155 GATCTCTAGTATTTCCCAAGGCAATATTAAGGTAGATAGATAGTATAGTATATTTGT 214  
QY 2766 TACACACTATTTTGGAAATAGAGAAATACAGAGAAATTTAGGGCTTAAACATTACGA 2825







```

RESULT 15
US-10-198-846-8531/c.
; Sequence 8531, Application US/10198846
; Publication No. US2003009974A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSIT
; TITLE OF INVENTION: FOR IDENTIFICATION
; TITLE OF INVENTION: THERAPY OF BREAST CA
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ IDS: 14084
; SOFTWARE: FASTQ for Windows Version 4.0

```

RESULT 14  
US-10-040-862-2900  
Sequence 2900: Application US/10040862  
Publication No. US20030078396A1  
GENERAL INFORMATION:  
APPLICANT: Galger, Alexander  
APPLICANT: Algate, Paul A.  
APPLICANT: Mannion, Jane  
APPLICANT: Retter, Marc  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
of Hematological Malignancies  
TITLE OF INVENTION: Hematological Malignancies  
FILE REFERENCE: 014058-013520US  
CURRENT APPLICATION NUMBER: US/10/040.862  
CURRENT FILING DATE: 2001-11-06  
PRIOR APPLICATION NUMBER: US 60/186,126  
PRIOR FILING DATE: 2000-03-01  
PRIOR APPLICATION NUMBER: US 60/190,479  
PRIOR FILING DATE: 2000-03-17  
PRIOR APPLICATION NUMBER: US 60/200,545